UNIVERSITY
OF MICHIGAN
DEC 13 1954
SOCIAL SCIENCE
LIBRARY

Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
AIR FORCE PERSONNEL AND TRAINING RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *AF Personnel and Training Research Center*
W. R. Garner, *The Johns Hopkins University*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *AF Personnel and Training Research Center*
Arthur L. Irion, *Tulane University*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

Identification of Visual Patterns as a Function of Information Load: E. J. ARCHER 313
The Effect of Context Stimuli on Learning and Retention: W. WEISS AND G. MARGOLIUS .. 318
Probability of Response to Compounds of Discriminated Stimuli: M. Ş. SCHOEFFLER 323
Intertrial Association at the Visual Threshold as a Function of Intertrial Interval: G. COLLIER 330
Acquisition of a 24-Unit Verbal Maze as a Function of Number of Alternate Choices Per Unit: W. J. BROGDEN AND R. E. SCHMIDT 335
A Further Investigation of Response Selection in Simultaneous and Successive Discrimination: A. D. CALVIN AND J. L. SEIBEL 339
Generalization of Expectancy Changes as a Function of the Nature of Reinforcement: H. M. SCHRODER AND J. B. ROTTER 343
The Role of Drive Reduction in the Classical Conditioning of an Autonomically Mediated Response: D. ZEAMAN AND N. WEGNER 349
Influence of Awareness of Reinforcement on Verbal Conditioning: J. B. SIDOWSKI 355
The Relationship Under Stress Between Changes in Skin Temperature, Electrical Skin Resistance, and Pulse Rate: L. M. BAKER AND W. M. TAYLOR 361
The Perception of the Egocentric Orientation of a Line: I. ROCK 367
The Effect of Component Practice on Performance of a Lever-Positioning Skill: G. E. BRIGGS AND W. J. BROGDEN 375
Multiple Response Transfer as a Function of Supplementary Training with Verbal Schematic Aids: F. H. KRESSE, R. M. PETERSON, AND D. A. GRANT 381
Two- and Three-Dimensional Autokinetic Movement as a Function of Size and Brightness of Stimuli: W. EDWARDS 391
Differential Conditioning and Level of Anxiety: K. W. SPENCE AND R. S. BEECROFT 399
The Relation of Electric Shock and Anxiety to Level of Performance in Eyelid Conditioning: K. W. SPENCE, I. E. FARBER, AND E. TAYLOR 404

American Psychological Association

Vol. 48 No. 5 November 1954

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1954 the publication cost is $20 per printed page. Authors of priority publications receive no free offprints.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Air Force Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail.

To save printing costs supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 [4, Suppl.], 389–449). In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 25th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Copyright, 1954, by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 48, No. 5 NOVEMBER, 1954

IDENTIFICATION OF VISUAL PATTERNS AS A FUNCTION OF INFORMATION LOAD[1]

E. JAMES ARCHER

University of Wisconsin

Much of behavior is directed by the information an organism receives. Quite often more information is received than can be effectively used, either because the organism cannot assimilate all of the relevant information or because some of it is irrelevant to the organism's behavior. In a simple example, a traffic sign bearing the word "Stop" transmits a certain amount of information which is relevant to the behavior of an automobile driver. The size, color, brightness, location, and even shape are irrelevant. Except for their constancy and consequent redundancy these last named attributes of the sign do not increase the presented information although they may increase the amount transmitted. When it is redundant, irrelevant information may increase information transmission, but when not redundant, information transmission will not be increased and may even be degraded. A study by Miller (4) which compared simple and complex conditional stimuli suggests that redundancy may increase information transmission even in a relatively simple communicative event. More conditioning occurred with two conditional stimuli than with either alone. From simple discrimination problems to concept-formation tasks, one of the primary needs of *S* is to identify the relevant information provided by the stimulus. The amounts of relevant and irrelevant information in a stimulus may vary independently of each other; task complexity is a function of amount of either or both types of information.

In a simple discrimination–reaction-time problem, Gregg (3) has shown that reaction time to visual stimuli increases as amount of irrelevant information increases. In his study, *S*s were required to move a joy stick either left or right depending upon the position, size, or brightness of a single dot on a dark field. Reaction time was longer as the irrelevant information was increased.

In a slightly more complex task, Archer, Bourne, and Brown (1) have found that time to solve a concept-formation problem increased as amount of irrelevant information increased. In this case, however, it may be suspected that *S* would attempt to incorporate the irrelevant information into hy-

[1] This paper is a modification of a Technical Report written on Contract AF 18(600)-54 between the University of Wisconsin and the Psychology Branch, Aero-Medical Laboratory, Wright Air Development Center, Wright Patterson Air Force Base, Dayton, Ohio.

potheses about the proper solution of the task and thereby be delayed since, by definition, irrelevant information could not enter into the correct solution of the concepts.

In Gregg's study, *S* had but one response to make and this was to one bit of relevant information. The effect of increasing amounts of irrelevant information was nonlinear because of the factor of stimulus-response compatibility.

In the Archer, Bourne, and Brown study the degradation of response time was nonlinear since each additional bit of irrelevant information doubled the number of necessary tests of hypotheses. It is conceivable, and in fact likely in view of Gregg's results, that irrelevant information may slow down *S* without actually being tested in a hypothesis. Furthermore, it is possible that there would be an interaction between amount of relevant and irrelevant information. Two predictions which might be made are that the amount of irrelevant information not only increases reaction time but will show an even greater effect as the amount of relevant information increases.

Procedure

Subjects.—Seventy-two paid male students at the University of Wisconsin served as *Ss*. Each *S* was assigned randomly to one of 12 conditions and served for one day for about 40 min.

Apparatus.—A device designated as a Target-Concept–Identification Apparatus was used. It consists of a cathode ray oscilloscope; *S*'s control unit, which has four lever-action switches and a push button; a pattern generator; and a recording console. The pattern generator, which is connected by electrical cables to the oscilloscope and *S*'s control unit, is operated by *E* through the use of two Western Union tape transmitters which program the patterns presented to *S*. The patterns are presented on the oscilloscope and may be classified according to one of two levels on each of six dimensions. These dimensions are form (circle or ellipse), size (small or large), brightness (bright or dim); speed (fast or slow), vertical direction (up or down), and horizontal direction (right or left). Since all patterns in this study travelled on a 45° or 315° diagonal, each had horizontal and vertical components. The *S*'s control unit had four lever-action switches; therefore only four of the six dimensions could be classified by *S*. The pattern generator had interchangeable leads which permitted a selection of the particular dimensions to be classified. The *S*'s control unit had interchangeable markers which signified the functional positions of the four lever-action switches. When the pattern generator was wired for a particular series of patterns, *S*'s control unit was modified accordingly.

The *Ss* saw patterns on an oscilloscope screen and classified these on four dimensions. Some dimensions were constant and contributed no information. Other dimensions varied between two levels, randomly and equally often. When one of the four classifying switches corresponded to this variable stimulus, the information was relevant. When a switch was not available, the information was irrelevant. The primary response measure was total time in minutes to classify the 32 patterns.

The recording console had two counters, a Standard Electric clock (.01-min. scale) and a program-starting button. After *S* was given a "ready" signal, the program-starting button was pressed. This action simultaneously presented the first pattern to be classified and started the timer clock. One of the counters recorded correct responses and the other recorded errors. When 32 patterns had been classified, the clock automatically stopped and *S* received a signal to stop working. Each *S* was instructed in the operation of the control unit as well as how to identify the levels of all dimensions. When a pattern was presented, he was to position each of the four lever-action switches according to the levels of each of the four dimensions for which he had switches and then to press the push button. If, for example, the control unit had switches for speed, size, form, and horizontal direction and a small, bright, fast circle moved from upper left to lower right, *S* would have to identify the speed as fast or slow and move the switch with these markers to one or the other position. He would have to do the same for each of the other dimensions of size, form, and horizontal direction. After all switches had been positioned (they always returned to center after firing holding relays), *S* would press the push button. If his identification and positioning were correct, a green panel lamp would light. If he were wrong, a red panel lamp would light. In both cases, however, a new pattern was presented. Whenever the push button was fired, all holding relays were released thus requiring *S* to reposition all four switches again even if successive patterns were the same.

It was possible to have from zero to six of the dimensions vary within a particular series. Since each dimension could only vary between two levels, it was possible to present series having 0, 2, 4, 8, 16, 32, and 64 different patterns. The first of these, 0 different, is of no interest

to our problem. To repeat again the relationship between switches and information: if a dimension is not varied, no information is presented; if a dimension is varied and no switch for classifying it is available, then the information is irrelevant; if a dimension is varied and a switch is available, then the information presented is relevant.

Both levels of all dimensions were easily discriminated and *Ss* were instructed as to essential identifying cues.

Design.—A 3 × 4 factorial design was used. Three levels of irrelevant information were orthogonal to four levels of relevant information. Six *Ss* served under each of these 12 relevant-irrelevant conditions. Each *S* classified 32 patterns as quickly as possible and then had a 1-min. rest. This cycle was repeated six times for a total of 196 classifications. To reduce the likelihood of a particular combination of dimensions being unusually easy or difficult, a different combination of relevant and irrelevant dimensions was used for each *S*. This meant, for example, that in the two relevant-one irrelevant conditions, one *S* had size and brightness relevant and form irrelevant whereas another *S* had form and speed relevant and vertical direction irrelevant.

General.—All *Ss* were given tape-recorded instructions which stated that they had to position all four switches and press the push button as fast as possible. They were told to use one hand for the lever switches and one for the push button. Furthermore, they were instructed to never move more than one switch at a time and never to move a switch before a pattern was presented, no matter how unchanging the patterns might be. After the instructions all of the 64 possible patterns were presented until *S* could give evidence of verbally identifying each of the two levels of the six dimensions of every pattern. For some dimensions there was an absolute basis for judgment, i.e., form and direction. For brightness, speed, and size, since the two levels of these dimensions were relative, *S* had to memorize cues; e.g., the dim patterns were just barely visible whereas the bright ones caused the entire oscilloscope to glow with a green light. During the demonstration *S* was prevented from practicing moving the switches and the emphasis was placed on verbal identification.

The importance of speed was stressed, but *S* was also cautioned to get as many right as possible. He was further instructed as to the purpose of the red and green panel lamps in front of him as well as the cease-working signal—a bright white light which was lit during the 1-min. intertrial rest periods.

Results

Response time.—The dependent variable of major interest in this study was the time required by *S* to classify the series of 32 patterns. The mean response times for each of the 12 conditions of amount of relevant and irrelevant information are presented in Table 1. Although there are a few small reversals of trends within particular rows and columns, these were not significant.

TABLE 1

Mean Response Times (Minutes) for Each Condition

Bits of Irrelevant Information	Bits of Relevant Information				Mean
	1	2	3	4	
2	2.34	2.69	3.37	3.30	2.93
1	2.12	2.99	2.85	3.65	2.90
0	2.24	2.47	3.15	3.08	2.73
Mean	2.24	2.72	3.13	3.34	

An analysis of variance of these time scores is presented in Table 2. The differences among the row means of Table 1 (three levels of irrelevant information) were not significant, but the differences among the column means (four levels of relevant information) were highly significant ($F = 17.74$, with 3 and 60 *df*, $p < .001$). Since the amount of relevant information was a quantified and ordered variable, it was possible to

TABLE 2

Analysis of Variance of Response Times (Minutes) Required to Classify 32 Patterns

Source	*df*	Mean Square	*F*
Relevant	3	25.65	17.74*
Linear	1	74.97	51.70*
Quadratic	1	1.89	1.30
Cubic	1	0.08	
Irrelevant	2	1.59	1.10
Relevant × Irrelevant	6	2.30	1.59
Ss/Groups	60	1.45	9.28*
Trials	5	37.65	241.47*
Trials × Relevant	15	0.17	1.07
Trials × Irrelevant	10	0.29	1.84
Trials × Relevant × Irrelevant	30	0.16	1.01
Ss × Trials/Groups	300	0.16	
Total	431		

* $p < .001$.

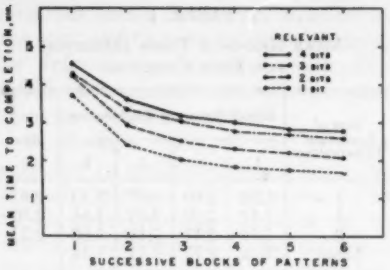

FIG. 1. Mean response time in minutes per block of 32 consecutive patterns as a function of stage of practice and amount of *relevant* information

further partition the sums of squares and associated degrees of freedom for the relevant information into the linear, quadratic, and cubic components through the use of orthogonal polynomials (2). As may be seen in Table 2 only the linear component was significant ($F = 51.70$, with 1 and 60 *df*, $p < .001$). This finding indicated that the time to respond increases as a linear function of amount of relevant information to be classified. The appropriate error term for evaluating the Relevant source and its three components is the *Ss*/Groups mean square. Bartlett's test indicated that the assumption of homogeneity of variance for this error term was tenable ($\chi^2 = 18.52$ with 11 *df*, $p > .05$).

The appropriate error term for evaluating individual differences (*Ss*/Groups) and Trials (and all interactions involving trials) is the *Ss* × Trials/Groups mean square. This term was composed of significantly heterogeneous variance estimates ($\chi^2 = 138.15$, 11 *df*), but no rational transformation, neither logarithms nor reciprocals, reduced the heterogeneity. In spite of this heterogeneity it appears reasonable to regard the differences among *Ss* and Trials as significant since the estimated p values exceeded the .001 level in both cases.

The effects of practice and amount of relevant information are shown in Fig. 1 for the four levels of relevant information, irrespective of amount of irrelevant information. The orderly increase in response time as a function of relevant information load is obvious.

An equation for response time as a function of information load, $Y_{min.} = .372X_{bits} + 1.925$, was obtained during the orthogonal polynomial analysis.

The effect of irrelevant information load on response time is shown in Fig. 2. Although the three curves may appear to be converging, neither the levels of irrelevant information nor the Levels × Trials interaction mean squares were significant. From Table 2 there was a suggestion that the Trials × Irrelevant Information interaction might be significant. An F of 1.836 was found, and for 10 and 300 *df* an F of 1.86 is required at the .05 level of confidence. Furthermore, from Fig. 2 it appears as if all three curves are converging and that there might be a difference in rate of classification as a function of irrelevant information only early in learning. An analysis of variance of time scores on the first trial indicated that this hypothesis was untenable ($F = 1.17$, $p > .05$, 2 and 60 *df*).

Errors.—Another response measure

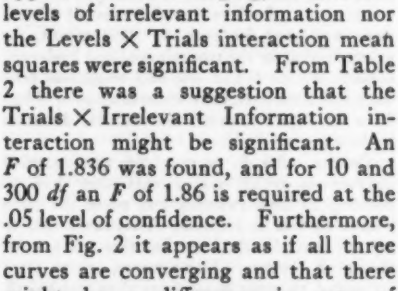

FIG. 2. Mean response time in minutes per block of 32 consecutive patterns as a function of stage of practice and amount of *irrelevant* information

which was obtained was number of errors made in responding to the 32 patterns. An error was automatically recorded if *S* pressed the push button before he had positioned all four switches or if he positioned one of them incorrectly. Even if more than one was incorrectly positioned, only one error was recorded. There was no systematic variation in errors except to decrease with trials ($F = 223.34$, $p < .001$, 5 and 300 *df*). No variation in number of errors appeared as either of the other two main effects of information load was manipulated.

Discussion

Although it may seem reasonable to assume that there were two components to the responses—(*a*) the motor response of moving a switch, and (*b*) the more mentalistic response of deciding in which position to move it—there was no evidence for this dichotomy; or at least if such did exist, practice facilitated both equally well. If this were not the case, the learning curves in Fig. 1 would not have been parallel.

It appears unlikely that the procedure used to measure reaction time in this study, total time to respond to a series of patterns, would account for the lack of agreement with Gregg's results (3). In his study the reaction time of each response was individually recorded. It seems more likely that the differences can be ascribed to the response requirements. Gregg's *S*s were first trained to respond to patterns which varied in one dimension only. In the second stage of the study, irrelevant information was introduced and the degradation of reaction time observed. The *S*s never had to make a response to more than one relevant dimension. In the present study, *S*s always had to respond to four dimensions of the stimulus, some of which varied. Unlike Gregg's *S*s, before the first trial *S*s also learned all of the possible patterns which could occur.

There are two findings of the present study which have significance for any theories of information transmission from a single stimulus source: (*a*) the amount of time required to decode and respond to information increases as a linear function of the amount of relevant information presented, at least up to four bits; and (*b*) response time appears to be unaffected by variations in information load as a function of amount of irrelevant information, at least up to two bits. The data also indicate that practice will improve performance in speed of pattern identification, but this effect is independent of amount of information load. Finally, practice will reduce the number of errors; however, this number is dependent upon neither the relevant nor irrelevant information present.

Summary

Twelve groups of six *S*s each served in an experiment on target identification. The groups corresponded to the cells of a 3×4 factorial design having one to four bits of relevant information and zero to two bits of irrelevant information, presented to *S* by a single stimulus source. The *S*'s task was to identify oscilloscope patterns by positioning four lever-action switches and to test this identification by pressing a push button. The response measure was the time required to identify 32 consecutive patterns. Errors were also recorded, but they did not vary as a function of any parameter except practice. The major findings were: (*a*) time to respond increases as a linear function of relevant information load, but (*b*) this response time was independent of amount of irrelevant information. These results were compared with those of previous studies.

References

1. Archer, E. J., Bourne, L. E., & Brown, F. G. Concept identification as a function of irrelevant information and instructions. *J. exp. Psychol.*, in press.
2. Fisher, R. A., & Yates, F. *Statistical tables for biological, agricultural, and medical research.* London: Oliver & Boyd, 1953.
3. Gregg, L. W. The effect of stimulus complexity on discriminative responses. *J. exp. Psychol.*, 1954, **48**, 289–297.
4. Miller, J. The rate of conditioning of human *S*s to single and multiple stimuli. *J. gen. Psychol.*, 1939, **20**, 399–408.

(Received March 16, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

THE EFFECT OF CONTEXT STIMULI ON LEARNING AND RETENTION[1]

WALTER WEISS AND GARRY MARGOLIUS

Boston University

The following statement is the basic theoretical proposition underlying the present research: when responses are made in the presence of perceptible stimuli, an association is developed between responses and the stimuli, such that the re-presentation of the stimuli will have a tendency to evoke the associated responses. Much of theory and research has been devoted to the development of this statement for the relationship between responses and, what might be termed, primary stimuli. A primary stimulus may be defined as the appropriate cue to which a response is directly associated. The stimulus syllable in a paired-associates experiment would be an example of this. It has been assumed by Guthrie (16), Hull (7), and McGeoch (8) that responses can become associated not only with primary stimuli but also with context stimuli. The color of the cards on which paired-associates are placed, general internal stimulation affecting the individual while he is responding, general stimulation derived from the environment in which the learning is occurring, would all be examples of context stimuli. Few studies, however, have been devoted to examining the role of such stimuli in *human* learning and retention.

On the basis of the assumption that such stimuli can become associated with responses being evoked, it seems appropriate to predict that the re-presentation of context stimuli, which had been present consistently while learning was occurring, would aid retention of the responses. Several investigations reported in the literature support this contention for some but not all of the context variables manipulated (1, 11, 13). A study by Pessin (12), however, did not confirm the hypothesis. The most carefully conceived and conducted research was one by Dulsky (2). Varying the color of the card on which paired-associates were placed, he found that recall is reduced when the color background present in the learning situation is changed for the retention test. (For the most part, available data, though not consistent or conclusive, indicate that under certain conditions, certain context stimuli provide support for the evocation of the desired responses. However, there are no data which indicate whether or not this support is maintained when the primary stimulus to which the response is associated is changed in the retention situation, i.e., when a generalized primary stimulus is presented. A major objective of the present study was therefore to check the hypothesis concerning the role in retention of context stimuli, using a procedure somewhat different from that of Dulsky, and to determine the effectiveness of the retentive support provided by context when the primary stimuli were changed.

A second purpose of the research was to determine whether or not the presence of a distinct, nondistracting context stimulus would affect the learning of a response. This consideration is different from the above-stated objective, in that the latter is

[1] This research was supported by funds made available by the Division of Research, Boston University, whose chairman is Dr. N. Maccoby.

concerned with the relationship between environmental stimulation in the learning *and* retention situations. The question now being raised is: regardless of the condition under which retention will occur, might there be any advantage for learning if certain context stimuli are presented throughout the learning session? Since the learning of complex responses requires a series of trials, it is conceivable, in accord with behavior theory (7), that the presence of an extraneous stimulus to which a desired response would become associated could provoke the arousal of that response in the early trials, before it is strongly associated with the appropriate primary stimulus. Learning, therefore, as measured by the presence or absence of the correct response where called for, would be aided.

Method

General design.—Four main groups of *Ss* were required to learn a list of nine paired-associates to a criterion of one errorless trial. Each pair of associates was placed on a solid-color card; a different color was used for each pair. Twenty-four hours following acquisition, all *Ss* were tested for retention and relearning under four different conditions. For Group 1, both color background and paired-associates were exactly the same as during acquisition. Group 2 had the original list, but now on uniform gray cards. Group 3 was presented with a generalized list, containing stimulus members that had been changed slightly from those used in acquisition; the response member of each pair was not changed nor were the colors of the cards changed. Group 4 relearned with the generalized list on homogeneous gray cards. Thus, Group 1 simply relearned under the original conditions; while Groups 2–4 relearned with certain variations in stimuli and/or the colors of the cards.

Two other groups were employed as controls. Group 5 was treated the same as Group 1 except for one variation in the condition of relearning: the stimulus member of each pair was not presented. This group provides a clearer view of the supportive utilization of contextual stimuli for retention and relearning. Group 6 learned the list on homogeneous gray cards and relearned under the same conditions as Group 5—colored cards and without the presentation of the stimulus members of the pairs. This group permits a comparison of the effect of colored versus gray backgrounds during acquisition. Also, the first-trial performance of this group is viewed as establishing a base line for guessing, against which to compare the performance of Group 5.

Apparatus and material.—The cards on which the paired-associates were typed were presented individually in a window cut in the center of a wide masonite screen. The stimulus member of each pair appeared in the left-hand part of the window; the response member in the right-hand part of the window. The *S* viewed the stimulus element alone for 3 sec. and then a shutter was raised and both stimulus and response elements appeared together for 3 sec. The cards were placed in a holding rack which could be slid horizontally along a track in back of the screen and under the window. As the cards could readily be removed from the rack, *E* was able to randomize the order and to remove specific cards at any time.

The nonsense syllables were selected from the Glaze list, as modified by Hilgard (6), and were of 80% association value. The stimulus element of each paired-associate was a hyphenated nonsense syllable; and the response element was a four-letter word. For example, two of the associates were MOT-FUD—RING and BIF-JOS—COAT. The generalized list was prepared by changing the first two letters of the second syllable, for example, MOT-WID—RING and BIF-NES—COAT (17).

Subjects.—The *Ss* were 48 volunteers from the general psychology course at Boston University. They were divided randomly into six equal groups.

Method.—After being shown two sample paired-associates within the apparatus, *S* was instructed to pronounce aloud the paired-associates of the entire nine-unit list without attempting to anticipate the response word before the shutter was raised. Three such passive viewings and vocalizations of the list were given before *S* was instructed to anticipate the response words. In accordance with the "adjusted-learning method" (9, 10, 15), if *S* correctly anticipated a response word, that card was removed from the rack. After all paired-associates of the list were shown, those cards incorrectly anticipated were shuffled and returned to the rack. Once again those correctly anticipated were removed and the remaining cards reshuffled for the next subtrial. When all paired-associates were correctly anticipated, and so removed from the rack, a single trial was considered complete. (A trial in this operational sense is the basis of the data analyzed in the section on Results.) All cards were then re-

turned to the rack and the next complete trial was run. This procedure was continued until *S* successfully anticipated all nine paired-associates during one presentation. This procedure assures equal correct anticipations for all members of the list, although the number of presentations for various paired-associates may vary.

The procedure during the relearning sessions was similar to that employed during acquisition except that *S*s were instructed to anticipate on the very first trial. The *S*s in Groups 1–5 were informed that the stimuli and color backgrounds might be altered somewhat. They were, however, to anticipate with the word they felt to be most appropriate. On the first relearning trial, *S*s in Group 6 who had no basis for correct anticipations were instructed to guess, utilizing the response words learned during the acquisition trials. Relearning continued until the criterion of one errorless trial was attained.

Results

Table 1 presents the learning and retention data for all groups. An analysis of the data on the acquisition of the stimulus-response pairs indicates no significant differences among the means of the groups that learned on colored cards. Thus, obtained differences in retention and relearning scores among these groups are not confounded with differing levels of original learning. When Groups 1–5 are combined into one aggregate unit, their composite mean number of trials to learn (5.02) is significantly lower than that for Group 6, which learned on uniform gray cards ($p < .02$, two tails). Acquisition of the primary stimulus-response associations is apparently facilitated as a result of the presentation of a particular color background with each paired-associate. An additional analysis made in terms of the mean number of total exposures of the paired-associates (i.e., the total number of presentations of the pairs) per group during acquisition supports the above findings.

An analysis of the recall scores, as determined by correct first-trial anticipations after 24 hr., for Groups 1–4 was made. All possible comparisons between the four groups are statistically significant, with p values ranging from $< .001$ to $< .05$, two tails. The main variables of color background and stimulus syllable have highly significant effects; but no interaction is found between them. The recall mean for Group 5 is significantly different from that for Group 6. This datum, together with the very close correspondence between the score of Group 5 and that of Group 3, which was given the generalized primary stimuli on colored cards, presents further evidence of the strong retentive support provided by the colored background.

For Groups 1–4 the main effect of color background, but not that of stimulus syllable, is found to be significant throughout relearning. Reliable differences related to the effect of stimulus syllable were not found on any trial subsequent to the initial recall trial. The relearning

TABLE 1

Learning and Retention Data for All Groups

Group	Stimulus Syllables During Retention	Background		Trials to Learn		Recall Scores		Relearning Trials	
		Learning	Retention	Mean	*SD*	Mean	*SD*	Mean	*SD*
1	Same	Colored	Colored	4.87	2.23	7.37	1.06	2.38	0.74
2	Same	Colored	Gray	5.12	2.04	3.25	2.55	5.25	1.83
3	Generalized	Colored	Colored	5.00	1.78	5.50	1.31	3.25	1.04
4	Generalized	Colored	Gray	4.75	2.07	1.37	1.19	4.88	1.55
5	None	Colored	Colored	5.37	1.19	5.37	1.19	2.75	0.89
6	None	Gray	Colored	6.62	1.58	0.75	0.89	5.38	2.67

score for Group 5 is significantly lower than that for Group 6 but is not reliably different from those for Groups 1 and 3, which also relearned on colored cards but with the stimulus syllables present. An analysis of mean number of errors during relearning confirmed and supported these findings.

Discussion

The research confirms the theoretical prediction that the representation of context stimuli, which had been present consistently while learning was occurring, would aid retention of the learned responses. The retentive support provided by associated context stimuli occurs not only when the original primary stimuli are present, but also when generalized primary ones are introduced. Thus, the findings of Dulsky (2) are confirmed and extended.

The data also support the expectation that the presence of distinct context stimuli facilitates acquisition of the stimulus-response associations. The actual process by which context stimuli can facilitate learning is not evident from the research. Two possible explanations can be suggested. It may be that the associative strength developed between the responses and the colors when combined with that developed between the responses and the primary stimuli raises the response strength above the evocation threshold in the early trials (7). Ordinarily, the strength of the primary association might not be sufficient to bring about evocation of the correct response. As the trials proceed, the associative strength between responses and primary stimuli alone will exceed the evocation threshold. A second possible explanation presumes that distinctly discernible context stimuli may act to reduce intralist generalization and thereby facilitate learning (5). Both of these processes may, of course, be occurring.

The effect of context stimuli on the retention of associations between responses and their appropriate stimuli depends on several considerations. Primary is the necessity for the development of an association between the correct response and the context stimuli. The conditions under which this occurs as well as the number of trials needed have not been established. Ferster (3) and Schoenfeld, Antonitis, and Bersh (14) indicate that the mere conjunction of response and stimulus may not be adequate. Fink and Patton (4), however, found that evocation of the correct response in the presence of environmental stimuli is a sufficient condition for the development of an association. The latter contention formed the basis for the current experimental procedure.

The nature of the context stimuli is also a factor. At the very least, they should not be of a kind to evoke responses interfering with the learning of the correct responses (i.e., distracting stimuli). Furthermore, they should be perceptually distinct and discriminable. Clearly different colors were employed in the present study to achieve favorable conditions for the demonstration of the predicted influence of context. That the colors proved to be so potent was somewhat unexpected. Whether or not context based on different stimuli would also prove to be as effective is not known. Nor is it established what the effectiveness of context would be if the primary and context stimuli form an integrated pattern of stimulation.

The effective influence of context stimuli does not depend alone on the strength of association between responses and such stimuli. When the effective association between responses and their primary appropriate stimuli is great, the presence or absence of contextual supporting stimuli need have no effect on response evocation in the retention situation. The presence of the primary stimuli would be sufficient. Thus, the choice of the number of trials in which the response is made to the primary stimuli would be critical for context-cue experiments. Too few trials would prevent the development of an adequate supporting association with context stimuli, while too many would develop so strong an association between responses and primary stimuli that the contribution of context stimuli would not

be detectible. (It is also to be expected that the optimum number of trials to show the contribution of context would vary with the type of learning task and the type of stimulation employed.)

The reduction of the strength of primary association by forgetting and by the use of generalized primary stimuli, as in the present research, represents two procedures that favor the demonstration of the role of context stimuli in retention. A third method not employed here would be to introduce new stimuli in the retention situation that have a tendency to evoke interfering, competing responses.

Summary

The *S*s learned paired-associates on differently colored cards. Recall and relearning of the associations were measured 24 hr. later under four conditions: same stimulus-response pairs on the same colored cards; same stimulus-response pairs on uniform gray cards; same response syllables but with generalized stimulus syllables on the same colored cards; same response syllables but with generalized stimulus syllables on gray cards. Two control groups were also employed: one learned the paired-associates on colored cards and was tested for retention and relearning with colored cards lacking the stimulus members of the pairs; the other learned the associates on uniform gray cards, and was tested for retention and relearning with the colored cards lacking the stimulus elements.

The data indicated that retention was better when the color backgrounds remained the same than when they were changed, whether or not changes were made in the stimulus syllables. The order of the four main groups in terms of mean retention scores from best to poorest was: (*a*) no change in stimulus syllable, no change in color background; (*b*) change in stimulus syllable, no change in color background; (*c*) no change in stimulus syllable, change in color background; (*d*) change in both stimulus syllable and color background. Relearning was found to be facilitated by the maintenance of the same colored backgrounds present during original learning. Also, the acquisition of the stimulus-response associations was found to be significantly better when colored contexts were employed than when a uniform gray background was presented. Some theoretical considerations with respect to context-cue experiments were noted.

References

1. Abernethy, E. M. The effect of changed environmental conditions upon the results of college examinations. *J. Psychol.*, 1940, **10**, 293–301.
2. Dulsky, S. G. The effect of a change of background on recall and relearning. *J. exp. Psychol.*, 1935, **18**, 725–740.
3. Ferster, C. B. The effect on extinction responding of stimuli continuously present during conditioning. *J. exp. Psychol.*, 1951, **42**, 443–449.
4. Fink, J. B., & Patton, R. M. Decrement of a learned drinking response accompanying changes in several stimulus characteristics. *J. comp. physiol. Psychol.*, 1953, **46**, 23–27.
5. Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.
6. Hilgard, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 517–567.
7. Hull, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
8. McGeoch, J. A. *The psychology of human learning*. New York: Longmans, Green, 1942.
9. Madden, M. S., Adams, J. A., & Spence, S. A. Memory-drum versus adjusted-learning techniques in the study of associative interference in learning by paired-associates. *Amer. J. Psychol.*, 1950, **63**, 186–195.
10. Osgood, C. E. Meaningful similarity and interference in learning. *J. exp. Psychol.*, 1946, **36**, 277–301.
11. Pan, S. Influence of context upon learning and recall. *J. exp. Psychol.*, 1926, **9**, 468–491.
12. Pessin, J. Effect of similar and dissimilar conditions upon learning and relearning. *J. exp. Psychol.*, 1932, **15**, 427–435.
13. Reed, H. J. The influence of a change of conditions upon the amount recalled. *J. exp. Psychol.*, 1931, **14**, 632–649.
14. Schoenfeld, W. N., Antonitis, J. J., & Bersh, P. J. A preliminary study of training conditions necessary for secondary reinforcement. *J. exp. Psychol.*, 1950, **40**, 40–45.
15. Shephard, A. H., & Lewis, D. A new device for presenting paired-associates in verbal-learning. *Amer. J. Psychol.*, 1950, **63**, 600–606.
16. Smith, S., & Guthrie, E. R. *General psychology*. New York: Appleton, 1921.
17. Yum, K. S. An experimental test of the law of assimilation. *J. exp. Psychol.*, 1931, **14**, 68–82.

(Received March 11, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

PROBABILITY OF RESPONSE TO COMPOUNDS OF DISCRIMINATED STIMULI [1]

MAX S. SCHOEFFLER

Indiana University

This study is designed to determine the probability of a response when the stimulating situation that is presented is composed in part of a stimulus to which that response has been conditioned, and in part of a stimulus to which a competing response has been conditioned. The experiment goes beyond a related study by Gulliksen (2) in that it utilizes training and test conditions which can be readily related to a theory which provides quantitative predictions regarding response probability in such a compounded stimulus situation.

In the present study, a group of 24 lights was randomly divided for each *O* into three distinct groups of eight lights each (S_I, S_{II}, S_{III}). A group of *O*s was conditioned to respond by moving a lever in one direction when S_I was presented, and by moving it in the other direction when S_{II} was presented. When *O*s had learned to make the correct responses with a high probability of success, a trial was presented on which predetermined proportions of lights from S_I, S_{II}, and S_{III} were presented simultaneously. The proportion of *O*s responding to the side that had been associated with S_I during the discrimination training was taken as a measure of the probability of a response in the S_I direction under this compound stimulation.

[1] This study was taken from a dissertation submitted to the Department of Psychology of Indiana University in partial fulfillment of the requirements for the Ph.D. degree in psychology. The writer is greatly indebted to Professor W. K. Estes under whose direction this investigation was conducted.

The set theoretical model of learning developed by Estes and Burke (1) provides quantitative predictions regarding the probability of a response in the S_I direction. In this theory, the sets of lights, S_I, S_{II}, and S_{III}, are conceived of as nonoverlapping stimulus situations, each of which consists of a large number of stimulus elements. If, for example, the lights comprising set S_I are lighted on a given trial, it is assumed that a random sample of the elements associated with these lights is presented to *O*. Each light is assumed to have the same number of elements associated with it, and to each element there is associated a probability that it will be sampled when its light is presented. For purposes of simplification, it can be assumed that all elements have the same probability (θ) of being sampled when their corresponding lights are presented. The parameter θ, then, also represents the proportion of elements in a set that are sampled when the set is presented. These elements can become conditioned to a response, and under the conditions assumed in the present study, the probability of a response is equal to the proportion of elements in a set which are conditioned to that response. The *O*s learn the discrimination until they reach a high probability of success. If it is assumed that the discrimination is perfect, all the elements in S_I are considered to be conditioned to a given movement of the lever, and all the elements in S_{II} to the opposing response.

Prediction as to the probability of a

movement in the direction associated with S_I, when lights from S_I, S_{II}, and S_{III} appear simultaneously, follows directly. If, for example, all the lights from S_I appear together with half of the lights from S_{II} on a test trial, it is possible to make the following analysis. Since all the elements associated with the S_I lights are conditioned to a certain response, the quantity $N\theta$, of conditioned elements is present, N being the total number of elements in any one of the sets. But at the same time half of the elements from S_{II} appear and none of these elements are conditioned to this response since they all have been conditioned to the opposing response. The proportion of the elements appearing that are conditioned to the response in the S_I direction is then $\frac{N\theta}{N\theta + \frac{N\theta}{2}}$, and the probability of this response is $\frac{2}{3}$.

This theoretically derived probability can be quantitatively compared with the observed proportion of *O*s responding to the side conditioned to S_I.

In a similar way, exact predictions can be made for other proportions of lights from S_I, S_{II}, and S_{III} presented simultaneously. If the lights from S_{III} are presented on a trial, it is assumed that since these lights have never appeared previously, half the elements corresponding to these lights are connected to a left movement and the other half to a right movement—the only responses available to the *O*.

Method

Subjects.—The *O*s for this experiment were 234 men and women undergraduates at Indiana University. The majority of *O*s were obtained from introductory psychology classes on a promise of extra points in their course work. The remainder were recruited singly without the offer of a reward.

Apparatus.—Facing *O*, there was a panel of ¼-in. plywood, 24 in. wide and 23 in. high. It was painted black and contained four rows and six columns of holes behind each of which was mounted a 6-w. light bulb. The distance between centers of adjacent holes was 4 in., in both the horizontal and vertical directions. In order to prevent the light from shining through an adjacent hole, a 1-in. piece of aluminum tubing was inserted into each hole reducing the inside diameter of the hole to ¾ in. The tubing was flush with the panel on the side toward *O*, and a rubber washer with an inside diameter of $\frac{3}{16}$ in. was placed between the tube and the bulb on the other side. The rest of the bulb was then painted black so that light from the bulb could pass only through the hole in front of it. Centered directly above the panel were two additional lights, one marked R and one marked L.

In front of the panel toward *O*, there was a black wooden box. A ¼-in. brass rod projected from this box toward *O*. This lever could be moved either to the right or to the left until it actuated a microswitch. The lever was spring loaded by means of rubber bands so as to be self-centering, and it projected 12 in. from the panel. The lever was 4¾ in. above the table, and *O* was seated so as to be in a comfortable position to move the lever. There was a distance of approximately 3 ft. between the front of the panel and the eyes of *O*. The *E* sat on the opposite side of the panel from *O*, and a 30-w. fluorescent bulb was placed behind *E* to permit recording. The whole apparatus was mounted on a table in a sound- and lightproof room.

The lights on top of the panel facing *O* were arranged so that when *O* moved the lever on a training trial, one of the lights (either R or L) would go on, depending on which pattern of panel lights was presented. This light then remained on as long as *O* kept the lever to the side or until the timing device terminated the trial. On the test trials, neither the R nor the L light came on. Each trial was automatically timed to last for a period of 2 sec. after which the apparatus was inoperative for 4 sec., the time of a complete cycle being 6 sec.

Procedure.—At the beginning of a session, *O* was read the following set of instructions:

"The purpose of the experiment is to find out how people learn to react correctly to different sets of stimuli. In front of you there is a panel of lights. When the experiment starts, some of those lights will go on for a period of 2 sec. As soon as the lights go on you are to move that lever in front of you either to the right or to the left. As soon as you have moved it, one of the lights on top—the ones marked R and L—one of those two lights will go on, indicating which way you should have moved the lever. If you make a mistake, do not attempt to correct it,

but try to respond correctly on the next trial. You are to try to move the lever to the correct side on as many trials as possible—that is, to make as few errors as possible.

"On some trials neither light will go on so that in this case you will not know whether you have responded properly or not.

"Sometimes you may have difficulty trying to decide to which side to move the lever. If this happens make sure that you nevertheless move it during the 2 sec. in which the lights are on. If you have not moved before the lights go out again, it will count as an error, so even if you cannot make up your mind, move the lever immediately anyway, because you may have it correct.

"Are there any questions? Well, practice moving that lever a few times so that you get the feel of it. Well, we're ready to start now. Remember that you move the lever while the lights are on. Ready?"

If *O* had any questions, the appropriate parts of the instructions were reread until he stated that he had no further questions.

After these instructions had been read to *O*, S_I and S_{II} were presented in a random sequence (constructed separately for each *O*) with the restriction that within each block of ten trials there must be five S_I trials and five S_{II} trials. For half the *O*s, every time S_I was presented, the light marked L flashed when *O* responded and every time S_{II} was presented, the light marked R operated. For the other half of the *O*s, these conditions were reversed.

Thirty of these training trials were given to each *O*. If *O* made more than one error during the last ten trials he was dismissed. The *O*s who made either one or no errors during the last ten trials were continued in the experiment, and will hereafter be called "used" *O*s. On Trial 31, all these used *O*s were given the first of ten test trials. Between any two of these test trials some more training trials were given. The number of these varied randomly from three to five, with the restriction that for each *O*, three pairs of test trials were to be separated by three training trials, three pairs by four training trials, and three pairs by five training trials. In this way, all *O*s received the same total number of training trials (66 trials). There were nine different kinds of test patterns and each *O* received all of these. On the tenth test trial a given *O* again received the same condition he had had as his first test trial. This was done in order to permit analysis to determine the effect of additional training encountered when a test trial came late in the series. The nature of the test trials is described below.

The nine different kinds of test patterns corresponded to the varying proportions of lights from S_I, S_{II}, and S_{III} that were presented. A description of these trials is given in Table 1.

TABLE 1

THE PROPORTIONS OF LIGHTS FROM S_I, S_{II}, AND S_{III} PRESENTED ON A GIVEN KIND OF TEST TRIAL

Kind of Trial	Proportion			Total Number of Lights
	S_I	S_{II}	S_{III}	
1	1	1		16
2	1	$\frac{1}{2}$		12
3	1	$\frac{1}{4}$		10
4	$\frac{1}{2}$	$\frac{1}{4}$		6
5	1	$\frac{1}{2}$	1	20
6	1	$\frac{1}{4}$	1	18
7	$\frac{1}{2}$	$\frac{1}{4}$	1	14
8	1	0	1	16
9	1	1	1	24

As mentioned previously, the 24 lights were divided randomly for each *O* into three groups of eight lights each. In order to determine which lights would be included in a given test pattern, the necessary number of lights was chosen randomly from each of the previously determined sets of eight, separately for each *O* and each test condition.

In order to balance for practice effect resulting from the fact that each *O* received all test patterns, the used *O*s were divided into nine groups of 20 *O*s each, and the test patterns were presented in the same order for all *O*s within each group. Each group received the test patterns in the order prescribed by the randomly chosen latin square in Table 2. The rows represent the different groups of 20 *O*s; the columns represent the ordinal positions of the test trials; and the entries in the cells represent the types of test trials.

Each of the groups of 20 *O*s arrived at in this way was further subdivided into two groups of ten each, half of the *O*s being trained to move right for S_I and the other half being trained to move left for S_I. The *O*s were assigned to groups haphazardly as they appeared for their appointments, but no group was permitted at any time to have more than two used *O*s in excess of any other group. In this way *O*s were added to each group until all groups had their quotas of 20 used *O*s. Altogether, 234 *O*s took part in the experiment, including 54 who were discarded for not fulfilling the criterion of no more than one incorrect response in the last 10 of the first 30 trials.

RESULTS

Figure 1 shows the curve describing the learning of the discrimination

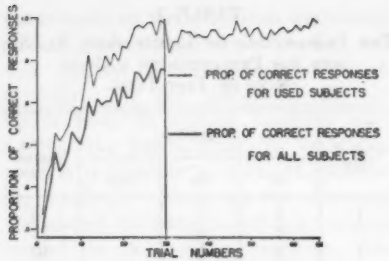
PROPORTION OF CORRECT RESPONSES
PROP. OF CORRECT RESPONSES FOR USED SUBJECTS
PROP. OF CORRECT RESPONSES FOR ALL SUBJECTS
TRIAL NUMBERS

FIG. 1. Discrimination learning curves for all *Os* and for "used" *Os*. Test trials appearing after 30th training trial in curve for "used" *Os* are not shown.

based upon the two nonoverlapping sets of lights. The proportion of correct responses on each of the 30 training trials is plotted for all *Os* who served in the experiment, including the *Os* who did not fulfill the criterion of at least nine out of the last ten responses being correct, and who were thereafter discarded from the rest of the experiment. Also shown is a curve constructed for only *Os* fulfilling the learning criterion; and this curve is extended for the complete set of 66 trials that these *Os* received. It must be noted here that the test trials which were given to the used *Os* after Trial 30 and at three to five trial intervals thereafter are not shown on this curve.

The number of *Os* responding to the S_I side on the test trials is given in Table 2. The test pattern associated with each cell is indicated by a roman numeral. The total possible entry in each cell is 20.

An analysis of this latin square indicates that there are significant differences among conditions but no significant differences with respect to groups or positions of test trials. The

TABLE 2

LATIN-SQUARE ARRANGEMENT SHOWING NUMBER OF RESPONSES TO S_I MADE BY EACH GROUP OF *Os* TO EACH TEST PATTERN

Groups of *Os* $N = 20$	Ordinal Position of Test Pattern								
	1	2	3	4	5	6	7	8	9
1	I 14	VIII 16	VI 11	V 15	IV 9	II 18	III 18	IX 8	VII 12
2	V 11	VI 11	III 18	I 9	VIII 11	IX 8	II 16	VII 11	IV 10
3	IX 11	I 9	II 15	IV 12	VII 13	III 17	VIII 13	VI 15	V 12
4	VIII 14	IV 12	VII 10	II 16	III 15	VI 14	I 9	V 13	IX 13
5	III 15	II 17	IX 16	VIII 14	VI 12	VII 12	V 13	IV 13	I 10
6	VI 14	V 15	IV 14	IX 9	I 9	VIII 16	VII 9	III 16	II 14
7	IV 14	III 16	VIII 17	VII 10	V 13	I 12	IX 10	II 14	VI 16
8	II 16	VII 10	V 9	III 16	IX 11	IV 11	VI 14	I 13	VIII 15
9	VII 11	IX 12	I 12	VI 14	II 16	V 11	IV 18	VIII 15	III 15

TABLE 3

LATIN-SQUARE ANALYSIS OF THE RESULTS GIVEN IN TABLE 2

Source	*df*	Sum of Squares	*F*
Groups	8	29	0.931
Positions	8	12	0.385
Conditions	8	316	10.146*
Residual (error)	56	218	

* $p < .01$.

results of the analysis are given in Table 3.

As an additional test to determine whether it is important at which point in the experiment a particular test pattern is given, each *O* was presented on the last trial of the experiment (the tenth test trial) with the same test pattern as had been presented to him on his first test trial. Thus the 20 *O*s in Group 1 were given Test Pattern 1 on their first test trial and again on their last test trial. Similarly, Group 2 *O*s were given Test Pattern 5 on their first and last test trials, and so on. The numbers of responses each group made to the S_I side when the repeated test pattern appeared in the first and in the last positions were then compared. These values, together with the result of a *t* test for paired measures, are given in Table 4. The value of $t(1.34)$ indicates that the position of the test trial in relation to the other test trials is not an important variable, a result which is in accord with the nonsignificant *F* for ordinal position obtained in the latin-square analysis.

TABLE 4

NUMBER OF RESPONSES TO THE S_I SIDE FOR THE REPEATED TEST PATTERN

Position	Test Pattern								
	1	2	3	4	5	6	7	8	9
First	14	16	15	14	11	14	11	14	11
Last	12	13	17	14	10	15	9	15	6

$t = 1.34$; $df = 8$

TABLE 5

COMPARISONS OF EXPECTED AND OBTAINED PROPORTIONS OF S_I RESPONSES TO THE TEST PATTERNS

Test Pattern	Expected Proportion to S_I Side	Obtained Proportion to S_I Side	*t*	*p*
1	.500	.539	1.05	.294
2	.667	.789	3.49	.0004
3	.800	.811	.37	.712
4	.667	.628	1.11	.266
5	.600	.622	.61	.542
6	.667	.672	.14	.888
7	.571	.544	.70	.484
8	.750	.728	.69	.490
9	.500	.544	1.19	.234

In order to test the theoretical predictions concerning compounding, a computation was made of the total proportion of *O*s who responded to the S_I side in the presence of each of the different test patterns. For this purpose the responses were combined for the various groups in the latin square. The proportions for the various test patterns were then compared with the proportions expected on theoretical grounds. The comparisons are given in Table 5.

These comparisons show that none of the differences between obtained and predicted values even approach significance except for the one for Test Pattern 2. Another test of

TABLE 6

COMPARISON OF EXPECTED ORDER OF GROUPS ACCORDING TO MAGNITUDE WITH OBTAINED ORDER

Test Pattern	Expected Proportion to S_I Side	Observed Proportion to S_I Side
3	.80	.811
8	.75	.728
6,4,2	.667	.696
5	.60	.622
7	.571	.544
9,1	.50	.542

agreement between theory and data was obtained by ordering the expected proportions with respect to size. This ordering was then compared with the ordering of the obtained proportions. For this purpose, the test conditions for which the same proportions were expected were averaged. An identical ordering of the two sets of numbers resulted, as may be seen in Table 6. The probability of this identity in order occurring on the basis of chance is one in 720.

Discussion

The experiment was in effect done in two stages. In the first stage, *Os* learned to respond differentially to two stimulus situations. After they had mastered this discrimination with a high probability of success, they were presented with different situations in which the stimulation was compounded of various proportions of the stimuli to which different responses had previously been conditioned. Thus, in one stage of the experiment, responses were measured in relation to isolated segments of a field of stimulation, and in the second stage, the segments of the field were combined in various ways to provide data concerning the effect of such combinations on the behavior of *O*, and to test in a quantitative manner some predictions which were derived from the set theoretical model of learning developed by Estes and Burke.

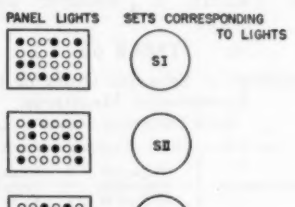

Fig. 2. Schematic diagram of disjoint sets of panel lights corresponding to the three stimulus sets S_I, S_{II}, and S_{III}

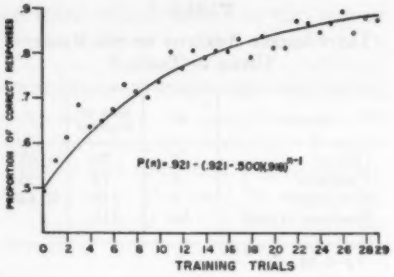

Fig. 3. Discrimination learning curve for all *Os* together with a theoretical curve obtained by minimizing the squared deviations

The two stages of the experiment and the results obtained in them will be discussed separately below.

Discrimination learning.—In the present experiment, the probability of occurrence of a response in the presence of a given stimulus, is related to the proportion of elements in the corresponding stimulus set that are connected to that response. There are only two responses available to *O*, and he is trained to make one of these consistently to a given stimulus set, and the other response to the other stimulus set. When this is achieved, all the elements in one stimulus set are considered to be conditioned to one response, and all the elements in the other stimulus set, to the other response. The schema diagrammed in Fig. 2 may clarify the conceptual interpretation of the experiment. The probability of a correct response on the nth trial $[p(n)]$ can be now described by an equation of the form: $p(n) = A - [A - p(1)]B^{n-1}$. The derivation of this equation is described elsewhere (3).

The trial-by-trial learning curve for all *Os*, shown in Fig. 1, is presented again in Fig. 3 together with a least-squares fit of the theoretical curve based on the above equation.

Compounding of stimuli.—In order to derive predictions from the Estes-Burke theory described earlier concerning the probability of response in the presence of compounded stimulation, certain assumptions are made, and the extent to which they are met must be considered.

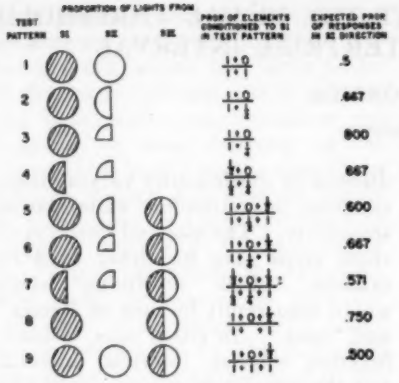

FIG. 4. Proportions of lights from each of the stimulus sets comprising the test patterns, together with probability of a response in the direction associated with S_I. Shaded areas represent elements in the set conditioned to a response in the S_I direction.

The assumption that the proportion of lights presented corresponds to the proportion of elements from the theoretical stimulus set seems reasonable, since the lights are separately manipulable, randomly drawn, and of equal wattage rating. Further, since only those *Os* which showed a well-learned discrimination were used in this part of the experiment, it can be reasonably assumed that all of the elements in S_I are conditioned to one response and all of the elements in S_{II} are conditioned to the other.

With the relation between the theory and the experiment thus clarified, and with other variables such as side preference and interactions among test conditions under experimental control, it is possible to derive predictions of response probability in the combined stimulation of S_I, S_{II}, and S_{III}. The way in which these predictions are made has been discussed in the introduction. Figure 4 shows these probabilities for each of the test patterns.

The close agreement between the expected and the observed results indicated in Tables 5 and 6 lends considerable support to the Estes-Burke set theoretical model of learning. It also indicates that the theory is fruitful for the enterprise of gaining information as to how to apply the results of the innumerable experiments which have been done in the artificially stable field where only one segment is allowed to vary, to situations which are more complex, i.e., where this stability does not obtain.

SUMMARY

The experiment was performed to determine the probability of a response in stimulus situations composed in part of stimuli to which this response had been conditioned, in part of stimuli to which an opposing response had been conditioned, and in part of stimuli to which neither response had been conditioned. The experiment also permitted a comparison between the obtained response probabilities and exact quantitative predictions of these probabilities derived from the set theoretical model of learning developed by Estes and Burke.

The subjects ($N = 234$) were trained to discriminate between two disjoint sets of lights, and were then given a series of test trials on which lights from the two discriminated sets, and lights that had not been presented previously, appeared simultaneously in varying proportions. The results indicated that: (*a*) the discrimination curve is a negatively accelerated, increasing function as predicted by theory; and (*b*) the probability of a response in the presence of a given test pattern is equal to the proportion of component stimuli that are conditioned to that response. Observed response proportions on test trials were in close agreement with the predictions derived from the set theoretical model.

REFERENCES

1. ESTES, W. K., & BURKE, C. J. A theory of stimulus variability in learning. *Psychol. Rev.*, 1953, **60**, 276–286.
2. GULLIKSEN, H. Transfer of response in human subjects. *J. exp. Psychol.*, 1932, **15**, 496–516.
3. SCHOEFFLER, M. S. Probability of response to compounds of discriminated stimuli. Unpublished doctor's dissertation, Indiana Univ., 1953.

(Received February 15, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

INTERTRIAL ASSOCIATION AT THE VISUAL THRESHOLD AS A FUNCTION OF INTERTRIAL INTERVAL [1]

GEORGE COLLIER

Duke University

Recent experiments (**1, 3, 4**) have shown that successive responses to a repeated, invariant, simple visual stimulus are not statistically independent. The probability of a "yes" (or a "no") response on any trial is associated with the response on the preceding trial or trials, resulting in runs of "yes" and "no" responses. Two kinds of association are possible. In the first kind the P_R (probability of response) shifts from trial to trial as a function of some factor independent of the outcome of the preceding trial(s), while in the second kind of association the shift in probability is some function of the outcome of the preceding trial(s). The present experiment attempts to decide between the two alternatives by determining the relation between amount of association and the length of the intertrial interval.

In order to produce intertrial association, a factor independent of the preceding response or responses would have to vary in some nonrandom fashion. Two alternatives present themselves. The first alternative is that the independent factor varies in a cyclic fashion, in which case the function relating the length of the intertrial interval and the measure of association would itself be cyclic. The second alternative is that the independent factor varies in a noncyclic fashion, that is, in a sequence of discrete or continuously varying steps resulting in nonrandom variations in sensitivity. The size and duration of these steps may be either fixed or random. Such sensitivity steps would also result in runs of "yeses" and "noes." In either case, a linear function relating intertrial interval and the association measure would be expected. This function would intersect the ordinate (association measure) at a finite point determined by the average size of sensitivity steps, and would intersect the abscissa (intertrial interval) at the point at which the intertrial interval was equal to the average duration of the steps. From this latter point on, no relation between the amount of association and intertrial interval would exist.

Three alternatives reasonably obtain for the response-outcome type association. The first alternative is that while the P_R is related to the outcome of the preceding trial(s), the effect of the preceding response is not affected by the lapse of time. In this case, no relation between intertrial interval and amount of association would be expected. The second alternative is that the increment or decrement in P_R added by the response on the preceding trial or trials to the P_R on a given trial is a decreasing function of the length of time between successive responses. In this case, the amount of intertrial association would be a decreasing function of the intertrial interval. If the momentary amount of association is directly proportional to the interval (or some function of the interval), the function

[1] The experiment reported in this paper was supported by a grant from the Duke University Council on Research. The author wishes to thank Dr. George Kish who participated in the preliminary studies on which the present experiment was based.

would intersect both the ordinate and the abscissa and no relationship would be expected beyond this latter point. If the momentary amount of association is inversely proportional to the interval (or some function of the interval), as would be the case if the amount of association decayed simply as a function of time, the function would be asymptotic to the ordinate, but would intersect the abscissa. The third alternative is that the increment or decrement in P_R added by the responses on the preceding trial is an increasing function of the increments in P_R on the preceding trials and is a decreasing function of the length of time between successive responses. If the momentary amount of association is directly proportional to the interval (or some function of the interval), the function would intersect the ordinate but would be asymptotic to the abscissa. If the momentary amount of association is inversely proportional to the interval (or some function of the interval), the function would be asymptotic to both axes.

Method

Apparatus.—The apparatus used in this experiment was a Model III Hecht-Shlaer Adaptometer. It presented, binocularly, a 3° circular patch 7° below a red fixation point at a viewing distance of 25 cm. for a duration of 0.2 sec. The luminance of the patch was continuously variable from 1.5 log μml to 5.4 log μml. The color of the patch was determined by a filter having a maximum transmittance of 480 mμ. Calibration was maintained photometrically. The flash duration was determined by a pendulum shutter. The eight intertrial intervals used, 1, 2, 3, 4, 5, 9, 12, and 30 sec., were determined by a synchronous motor operating the shutter trip. A head rest and a chin rest were used. The *S*s responded by use of a key. The apparatus was housed in a sound-deadened, light-tight, ventilated room.

Subjects.—Six paid university undergraduates, three men and three women, served as *S*s. All met or exceeded Naval visual standards as tested by the Orthorator (20/20 visual acuity, O. D., O. S.; no phorias; "color" normal). Three had extensive previous experience in visual experimentation, three did not. They were instructed to maintain fixation and to "push the lever to the right" if they saw the flash, and "to the left" if they did not. At the beginning of the experimental period each *S* had spent at least 20 min. in dark-adaptation glasses and 20 min. in complete darkness. The *S*s were run at the same time each day.

Procedure.—Each *S* received four preliminary days on a standard method of limits procedure (3). On the following seven experimental days, the six *S*s were run at a fixed luminance which was determined from their prior method of limits days so as to yield approximately 50% "yes" responses.

Seven intertrial intervals were used, 2, 3, 4, 5, 9, 12, and 30 sec., in this portion of the experiment. Six of the seven experimental days consisted of three blocks of 100 trials, each separated by 1- to 2-min. rest periods. One intertrial interval was assigned to each block. In order to hold the running time approximately constant, intertrial intervals 2, 4, and 12 (total time = 1800 sec.) and 3, 5, and 9 (total time = 1700 sec.) were assigned to alternate days. The order within days, between days, and between *S*s was counterbalanced over the six days. Thus, each *S* received three blocks of 100 stimulations at each of the six intertrial intervals, 2 to 12 sec. The first or the seventh day consisted of 100 stimulations at the 30-sec. interval.

One *S* was run for an additional four days, after two days of practice, on a 1-sec. intertrial interval. She received 6 blocks of 100 trials each day, making a total of 24 blocks for the four days.

Results

Probability of response.—The percentage of "yes" responses was calculated for each *S*, on each block of trials, and under each condition. The three values for each subject condition

TABLE 1

Analysis of Variance of the Probability of Response

Source	*df*	Mean Square	*F*
Interval	5	87.05	0.56
Block	2	322.23	2.07
Subject	5	4039.01	25.96***
I × P	10	297.48	1.87
I × S	25	116.85	0.73
P × S	10	92.80	0.58
I × P × S	50	159.10	
Total	107		

*** Significant at .1% level.

TABLE 2

z Scores

Ss	Intertrial Interval						
	2	3	4	5	9	12	30
1	3.79	2.56	3.31	2.36	0.77	0.49	−0.59
2	4.51	2.45	4.54	4.26	3.96	2.37	1.75
3	6.85	3.00	3.84	3.30	4.43	0.20	1.94
4	9.86	6.66	5.96	3.05	2.59	1.46	−0.36
5	4.32	3.57	3.06	3.24	0.88	1.50	3.05
6	1.87	0.24	1.88	0.14	1.28	−0.48	−1.07
Mean	5.20	3.08	3.77	2.73	2.32	0.93	0.79

on intertrial intervals 2–12 sec. were averaged, and an analysis of variance was performed on intertrial intervals, blocks, and Ss (Table 1). Only Ss proved to be a significant (.1% level) source of variation. Intertrial interval and the portion of the record, first, second, or third block, do not affect the probability of response. No one of the interactions approached significance.

Intertrial association.—As a measure of the intertrial association, a critical ratio (z) was used which was based on the comparison of the expected number of runs, for the obtained percentage of "yeses," with the obtained number of runs (2, p. 254). Since the number of trials was constant, this statistic yields a measure of amount as well as significance of association. It yields the same information as the ϕ calculated in the autocorrelation techniques; the relation between the two values is $2N\phi^2 = z^2$.

The three values for each subject condition were averaged. These values averaged over blocks are presented in Table 2. An analysis of variance, similar to the preceding one, was performed (Table 3). Of the main effects, intertrial intervals and subjects proved to be significant, at .1% level, while blocks again proved to be nonsignificant. The intervals by-blocks and the interval-by-subject-interactions proved to be significant at the 1% level, while the subjects-by-blocks interaction did not approach significance. The significant interval-by-subject interaction appears to be due to one S who showed a very small amount of association at the 2- and 3-sec. intervals, while the significant intervals-by-blocks interaction is due to the long intervals following a short interval having an increased amount of association, and the short intervals following a long interval having a reduced amount of association.

A graph of the relationship between the amount of intertrial association and the intertrial interval is presented in Fig. 1 (upper curve). As can be seen, the relation is a monotonically[2] decreasing, positively accelerated one which appears to be asymptotic to both axes. An omnibus chi-square

TABLE 3

Analysis of Variance of the Association Measure

Source	*df*	Mean Square	*F*
Interval	5	36.99	8.02***
Block	2	4.15	1.19
Subject	5	35.92	7.79***
I × B	10	7.41	3.96**
I × S	25	4.61	2.46**
P × S	10	3.48	1.86
I × P × S	50	1.87	
Total	107		

** Significant at 1% level.
*** Significant at .1% level.

[2] A possible inversion appears at 3 sec., although it does not reach significance.

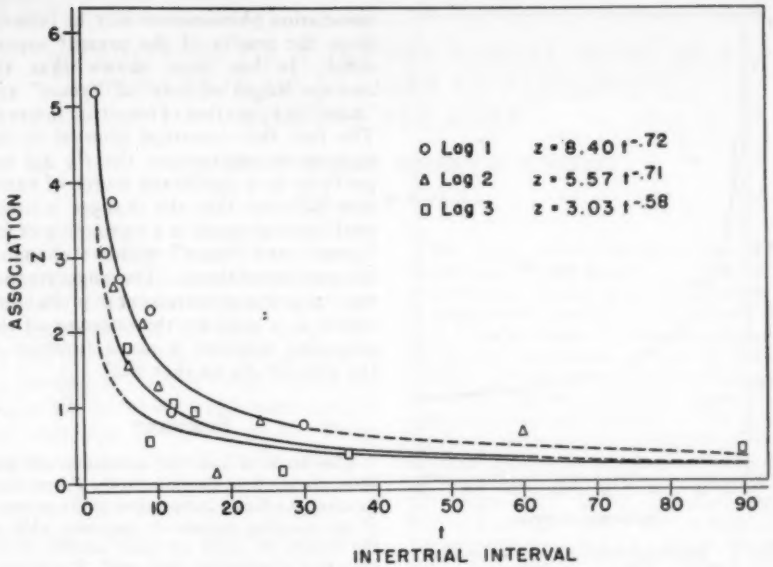

FIG. 1. Intertrial association as a function of intertrial interval. $N = 6$. The portion of curve for which data points existed was drawn as a solid line.

test shows that all values are significantly different, at the 1% level, from zero association, even at the 30-sec. intertrial interval.[3]

DISCUSSION

Of the six kinds of functions proposed, only the three seem plausible which hypothesize that the P_R shifts from trial to trial as a function of the outcome of the preceding trial(s) and that the amount of association is inversely related to the time between trials. Of these three functions, the one which assumes that the momentary amount of association is, further, related to the outcome of the preceding trials, rather than the just preceding trial, seems most likely, since the function appears to be asymptotic to both axes. On this hypothesis, a simple function of the form $z = at^{-b}$ was fitted by least squares to the data (Fig. 1, upper curve). A trend test gave no reason to reject the fit ($p > .2$; $F = 1.51$; df 4, 30). The correlation index between intertrial interval and amount of association was .624.

Additional support for the asymptotic character of the function was obtained from the results of one run at a 1-sec. intertrial interval (Fig. 2). These results indicate that the curve continues to rise rapidly as the intertrial interval decreases, although the point at 1 sec., which lies considerably above the predicted value, suggests that either the function chosen is too simple or that a further process intervenes at the short interval. A sampling error seems unlikely because of the large number of determinations made at this point.

Additional support for the hypothesis that the outcome of the preceding trials influences the momentary P_R was obtained by an analysis of the data for lag 2 and lag 3 (3). The amount of association was determined between every second response (lag 2) and every third response (lag 3). These values were averaged and plotted against the appropriate intertrial interval (twice the interval used for lag 2 and three times the interval for lag 3). Again the points appear to be along hyperboloid curves,

[3] The plot resembles, in form, the plot obtained in previous data for various lags (3).

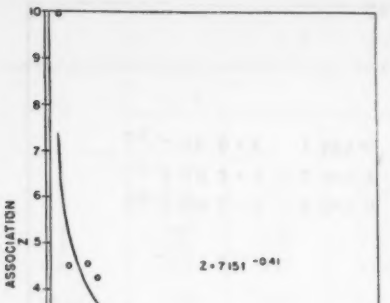

FIG. 2. Intertrial association as a function of intertrial interval. $N = 1$.

although the eccentricities of the lag 1 curve are increasingly exaggerated as the lag increases. The fitted curve for lag 2 lay below that for lag 1 (Fig. 1). The fitted curve for lag 3 was significantly below that for lag 1 ($p < .05$; $F = 4.77$; df 1, 30). If the intervening trials had no influence on the P_R, it would be expected, for example, that the point at intertrial interval of 3 sec. at lag 2 would fall at the same height as the point at 6 sec. on the lag 1 curve, since both trials have 6 sec. intervening between them. Only if the outcome of the intervening trials influenced the P_R would the curves for lag 2 and lag 3 lie below that for lag 1.

On the basis of the reasonable fit obtained, the conclusion of these arguments is, then, that the P_R under the present conditions of stimulation, is a function not only of the luminance used but also of the outcome of the preceding trials and that the effect of the preceding trials decays in time. It is obvious that this relationship will only hold over a finite range of values, since for the very short and very long intertrial intervals other processes, such as reaction time or motivation, must intervene.

One additional characteristic of the association phenomenon may be inferred from the results of the present experiment. It has been shown that the average length of runs[4] of "yeses" and "noes" is a function of intertrial interval. The fact that intertrial interval in the analysis of variance on the P_R did not prove to be a significant source of variation indicates that the changes in intertrial interval result in a regrouping of the "yeses" and "noes" without changing the number of them. This suggests that the increment or decrement in probability added to a trial by the outcome of the preceding response is not a function of the over-all P_R on that trial.

SUMMARY

Two kinds of intertrial association are possible. In the first, the shifts in the P_R are some function of a factor independent of the outcome of the preceding response or responses, while in the second, the shifts are some function of the preceding response or responses. The present experiment attempts to decide between these alternatives by determining the relationship between the amount of association and the intertrial interval.

A positively accelerated, decreasing function was obtained, which appeared to be asymptotic to both the association and time axes. Several possible functions were inferred from the two alternative types of association and, on the basis of the results, it was concluded that the intertrial association is a function of the outcome of the preceding trials, the effect decaying in time.

REFERENCES

1. COLLIER, G. Probability of response and intertrial association as functions of monocular and binocular stimulation. *J. exp. Psychol.*, 1954, **47**, 75–83.
2. DIXON, W. J., & MASSEY, F. J. *Introduction to statistical analysis.* New York: McGraw-Hill, 1951.
3. VERPLANCK, W. S., COLLIER, G. H., & COTTON, J. W. Nonindependence of successive responses in measurements of the visual threshold. *J. exp. Psychol.*, 1952, **44**, 273–282.
4. VERPLANCK, W. S., COTTON, J. W., & COLLIER, G. H. Previous training as a determinant of response dependency at threshold. *J. exp. Psychol.*, 1953, **46**, 10–14.

(Received January 6, 1954)

[4] The mean length run is equal to the number of responses divided by the number of runs.

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

ACQUISITION OF A 24-UNIT VERBAL MAZE AS A FUNCTION OF NUMBER OF ALTERNATE CHOICES PER UNIT[1]

W. J. BROGDEN AND ROBERT E. SCHMIDT

University of Wisconsin

The present study is an extension of an earlier experiment on the effect of number of alternate choices per unit on acquisition of a verbal maze (1). In the previous study, five groups of *Ss* learned a 16-unit verbal maze for which the number of alternate choices per unit was 2, 3, 4, 6, or 8. Total errors and total trials to the criterion of one perfect repetition increased linearly with increase in number of alternate choices. In discussing these results, it was hypothesized that the functions of acquisition and number of alternate choices per unit would become nonlinear at some greater level of maze difficulty, represented by a maze of greater length and with a larger number of alternate choices per unit. In the present experiment, *Ss* learned a 24-unit maze with 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 alternate choices per unit. Its purpose was to determine the relation between acquisition and number of alternate choices per unit at the greater level of difficulty represented by these mazes.

PROCEDURE

Ten *Ss* in each of 11 groups practiced a 24-unit verbal maze until the criterial trial of one perfect repetition was attained. The *Ss* were members of the classes in elementary psychology at the University of Wisconsin who volunteered to serve in an experiment on learning a mental maze. The *N* of 10 for each of the 11 groups was filled simultaneously by random assignment of *Ss* to each group.

The mazes were constructed from words representing angular spatial orientation in the vertical plane. Models of one unit of each of the 11 mazes were constructed from Tinker Toy parts. A shaft inserted in the center hole of a circular hub represented one unit of distance forward. The hub represented the choice point and shafts inserted around the circumference represented the alternate choices. The angular location of shafts and their designation in terms of verbal response are as follows: 0° = up (U); 30° = up-right (UR); 60° = right-up (RU); 90° = right (R); 120° = right-down (RD); 150° = down-right (DR); 180° = down (D); 210° = down-left (DL); 240° = left-down (LD); 270° = left (L); 300° = left-up (LU); 330° = up-left (UL). Cardboard tabs with the appropriate words printed on them were inserted at the end of each of the shafts attached to the hub circumference. The *E* used the models in giving instructions to *S*.

The sequence of correct choices making up each maze was obtained by use of a table of random numbers with the restrictions that the first choice be R and that there be an equal number for each kind of choice. Since the latter condition could not be met with the mazes for Groups 4, 6, 8, 9, and 10, equality for kind of choice was held to 4, 3, 2, 2, and 2, respectively; the correct choices for the remaining units were filled at random with the requirement that a given kind of choice not appear more than once. Where exact directional symmetry could not be met, two mazes of identical pattern were used and half the *Ss* of the group was run on each. The alternate choices for Group 1 were R and L. For Group 2, U or D was added to R and L. For Group 3, the choices were R, L, U, D. For Group 4, the addition of RU or RD was made. For Group 5, the choices were R, L, U, D, RU, RD. For Group 6, LU or LD was added. For Group 7, both LU and LD were added to the choices for Group 5. For Group 8, UR or UL was added. For Group 9, the choices were R, L, U, D, RU, RD, LU, LD, UR, UL. For Group 10, DR or DL was added. For Group 11, both DR and DL were added to the choices for Group 9.[2]

[1] This research was supported in part by a grant from the University Research Committee from funds provided by the Wisconsin Alumni Research Foundation.

[2] A table showing the exact pattern of each maze has been deposited with the American

The instructions are similar to those of the earlier study (1). They differ only in terms of the number and kinds of choices. The *S* was told the choices for his maze and the appropriate model was used in giving the instructions. When *S* indicated that he understood the instructions, he was blindfolded and the first trial began. The *E* recorded each error and the time of the trial on prepared data sheets. A 30-sec. intertrial interval was maintained throughout practice.

Results

An initial test of differences in the data for the two subgroups of Groups 2, 4, 5, 6, 8, and 10 was made. The mazes for these subgroups differed from each other only in terms of a few response words and not in terms of pattern or number of alternate choices per unit. Since no significant differences were found for either trials, time, or errors to the criterion, the data for the subgroups were pooled in each case. An analysis of variance was performed on the data for each of the three measures of acquisition. The *F* ratio for the between- and within-groups mean squares is 1.88 for trials to the criterion, 26.40 for time to the criterion, and 22.74 for errors to the criterion. The value of *F* at the 5% level of confidence is 1.92. The group means for trials to the criterion are 14.1, 20.8, 20.6, 22.3, 20.3, 21.5, 19.8, 20.6, 19.6, 20.4, and 19.5, respectively, for Groups 1 to 11. Except for Group 1, the means appear to be of comparable magnitude. The mean for Group 1 does not differ significantly from the means for the other groups.

Although the trial measures show no relation to number of alternate choices per unit, those for time and errors do. Both show a progressive increase in magnitude with increase in number of alternate choices per unit.

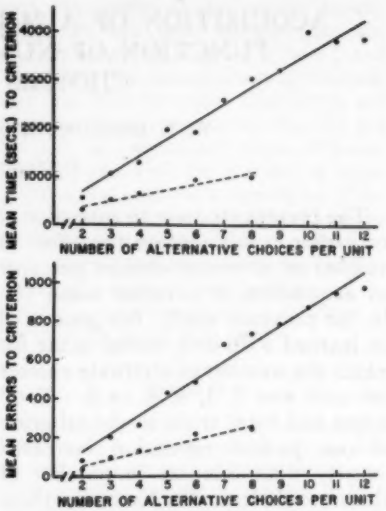

Fig. 1. Relation between acquisition of verbal mazes and number of alternate choices per unit

The solid dots of Fig. 1 are plots of these means. These relations were tested for deviation from linear regression (2). Since there was no significant deviation from linear regression in either case, straight-line regression equations were fitted to the data by the method of least squares. These equations are $y = 363.1x - 24.2$ for time and $y = 93.4x - 65.4$ for errors. The solid lines of Fig. 1 are plots of these curves. The open circles and the broken lines reproduce the results obtained with the 16-unit verbal maze already reported (1).

Total errors to the criterion were tabulated for each *S* for each of the 24 units in the maze and an analysis of variance performed on the data. The *F* for groups is 230.34, for units of the maze is 63.31, and for the interaction of groups by units is 8.62. Each of these *F* ratios is significant at better than the 5% level of confidence. The significant interaction of groups by units indicates that the form of the

Documentation Institute. Order · Document No. 4419 remitting in advance $1.25 for microfilm or $1.25 for photocopies.

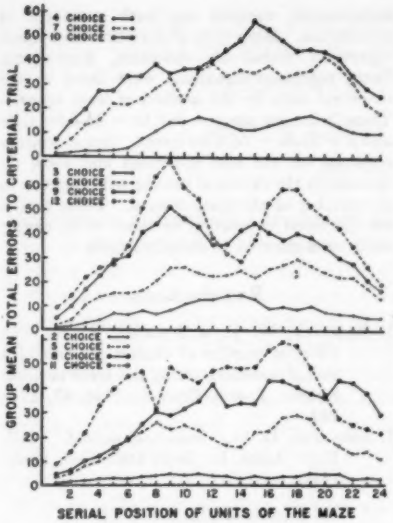

FIG. 2. Serial position error curves for a 24-unit verbal maze with different numbers of choices per unit

serial position error curves varies significantly between groups. The nature of this result is shown in Fig. 2 where the curves for each group are plotted. Each of the curves has been smoothed by taking a rolling average of "threes." Extent of bowing of the curves increases with increase in number of alternate choices per unit. For mazes with four and more choices per unit, there is a tendency for the curves to be bimodal, and it is clearly evident for the 12-choice maze. The first peak frequency occurs at about Unit 8–12 and a second peak frequency occurs at about Unit 15–18. Maximum frequency may occur at either of these two positions. For all curves, there is a greater frequency of error at the end of the maze than at the beginning.

DISCUSSION

The results of the present study are similar to those of the earlier study on the 16-unit maze (1). The results for trials may in each case be attributed to the correction procedure necessarily involved in acquisition of a maze-type task. Trials on mazes with different numbers of alternate choices are not comparable. The results for both errors and time show a progressive increase in magnitude with increase in the number of alternate choices per unit. For both the 16-unit maze of the earlier study and the 24-unit maze of the present study, these relations are linear (see Fig. 1). The difference in slope of the regression equations, being greater for the 24-unit maze, suggests that the slope of functions between time and error measures of acquisition and number of alternate choices will increase as the length of the maze is increased.

The results of the experiment have failed to provide any support for the hypothesis that the functions between time and error measures of acquisition and number of alternate choices per maze unit will become nonlinear at some greater level of maze difficulty, represented by a maze of greater length and with a larger number of alternate choices per unit than are involved in the 16-unit maze. They do indicate, however, that if such a hypothesis is correct, the shift from a linear to a nonlinear function will occur at a level of maze difficulty greater than that represented by the 24-unit length and 12 alternate choices per unit of the present experiment.

Discussion of the errors expected by chance on the first trial, errors actually obtained on the first trial, the factor of amount of discovery, and the relation of number of like responses in the correct pathway to number of alternate choices in the 16-unit maze is equally relevant to the present experiment. The interested reader is referred to the earlier paper (1). Note should be made, however, that with the 24-unit maze, repetitive errors on the first trial begin to occur on the maze with six alternate choices and, in general, increase in frequency as the number of alternate choices increases. This result is similar to that obtained with the 16-unit maze.

Similar results on the serial position

error curves hold for the two studies. Extent of bowing increases with increase in number of alternate choices. The curves are asymmetrical, there being a greater frequency of error at the end of the maze than at the beginning. The primary difference between the two experiments is the greater tendency for the curves for the 24-unit maze to be bimodal. This characteristic occurred only in the serial position error curve for the 8-choice, 16-unit maze of the earlier study.

Summary

Eleven groups of *Ss* learned a 24-unit verbal maze for which the number of alternate choices per unit were 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12. Acquisition measured in terms of mean total time and mean total errors to the criterion of one perfect repetition increased with increase in the number of alternate choices. Analysis of variance indicates a significant effect of the independent variable on both measures of acquisition. Since tests of deviation from linear regression reveal no deviation, least-square linear regression equations were fitted to the empirical data by the method of least squares. These functions are $y = 363.1x - 24.2$ for time and $y = 93.4x - 65.4$ for errors. Serial position error curves are also presented and show an increase in the extent of bowing with increase in the number of alternate choices. These results are discussed in relation to those of an earlier study on a maze of 16 units in length.

References

1. Brogden, W. J., & Schmidt, R. E. The effect of number of choices per unit of a verbal maze on learning and serial position errors. *J. exp. Psychol.*, 1954, **47**, 235–240.
2. Snedecor, G. W. *Statistical methods.* (4th Ed.) Ames, Ia.: Iowa State Coll. Press, 1948.

(Received March 3, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

A FURTHER INVESTIGATION OF RESPONSE SELECTION IN SIMULTANEOUS AND SUCCESSIVE DISCRIMINATION

ALLEN D. CALVIN AND JEAN L. SEIBEL

Michigan State College

Recently a great deal of discussion has appeared in the literature concerning the comparative difficulty of simultaneous and successive discrimination. The discussion was prompted by a study conducted by Weise and Bitterman (6). They used a multiple-unit alley maze with two small lamps located at each choice point. One group of rats was trained on a simultaneous problem where one lamp at each choice point was on and the other off. These *S*s had to learn to turn either in the direction of the light or dark alley at every choice point. A second group was trained on a successive problem where the lamps at each choice point were either both on or both off. These *S*s had to learn to turn in one direction when both lamps were on and in the opposite direction when both lamps were off. The successive group learned the problem much more easily than the simultaneous.

Weise and Bitterman felt that their results could not be handled by the theories of Nissen, Lashley, or Spence. The theoretical issues have been discussed at great length elsewhere and no attempt will be made in this paper to review the various arguments presented. However, aside from the theoretical points at issue, both Spence (5) and Nissen (4) raised objections to the design of the experiment. Nissen felt that the simultaneous group might have been hindered by reflected light. Spence felt that the complex type of discrimination apparatus used added problems such as the influence of the goal gradient. Bitterman and his associates have denied the validity of these claims (1).

Some experimental work with children by Calvin (2) led us to believe that Bitterman obtained his anomalous results primarily because of another factor in his experimental design. In the apparatus used by Weise and Bitterman the two lamps used as stimuli were presented so close together that they appeared as a single unified percept rather than as two discrete units. Such a situation would, of course, be favorable for the successive group, because it would tend to make it difficult for the simultaneous group to separate the positive and negative discriminanda. The study to be reported is a test of this hypothesis. The Weise-Bitterman experiment was repeated, except that the two lamps were moved out to the end of their respective maze arms, thereby breaking up the unity of the percept and creating two separate spatial units.

Meanwhile, Bitterman and Wodinsky (1) have recently presented an explanation quite similar to ours. They say: "In the apparatus of Weise and Bitterman, on the other hand, the two stimuli (lamps) at each choice point were closely juxtaposed and the animals were required to turn away from them, to one side or the other, in making their way through the maze. This may be described as a *response* situation. It facilitates configurational organization and thereby retards the functional isolation of the two members of each pair of stimuli. From this point of view the greater difficulty of the simultaneous problem is understandable either in terms of the greater similarity of its two configurations (on the assumption that solution is based on configurational discrimination) or in terms of the difficulty of perceptual analysis in such situations (on the assumption that solution is based ultimately on response to components). This interpretation, like Weise and Bitterman's, proposes two qualitatively distinct types of perceptual organization in these problems, but the functional priority of configurational perception is not postulated. Instead, it is assumed that the dominance of one or the other kind of organization is determined by the physical characteristics of the apparatus employed" (1, p. 373). In our apparatus, in which the lamps were 8 in. from the choice point at the end of their respective maze arms, the stimuli are not closely juxtaposed, and *S* had to approach the discriminanda; therefore, it seems apparent that Bitterman and Wodinsky would also predict

that the simultaneous group would be superior to the successive group.

Method

Apparatus.—Our apparatus duplicated Weise and Bitterman's except for the position of the lamps which served as the discriminanda, which were mounted 8 in. from the choice point. The starting box was 8 × 8 in., the food box 8 × 17 in., and there were four identical discrimination units. Each alley was 12 in. long and 4½ in. wide, and all the walls were 4½ in. high. The apparatus was painted flat black, and the top was covered with hardware cloth. For a detailed description of the equipment see Weise and Bitterman (6).

Procedure.—Our procedure was a duplicate of Weise and Bitterman's. We used 20 naive albino rats approximately 3 mo. old. Group I consisted of 6 males and 4 females and Group II of 5 males and 5 females, thus exactly replicating the make-up of their groups. Group I learned the simultaneous problem, and Group II learned the successive problem.

For the simultaneous group only one light at each choice point was on. Half the *S*s in Group I had the light positive and half had dark positive. The following six random orders utilized by Weise and Bitterman were used: RRLL, RLRL, RLLR, LLRR, LRLR, and LRRL. These were employed in a random fashion on successive trials.

For the successive group the lights at each choice point were either both on or both off. Half the *S*s had to go left to light and right to dark, and the other half had to go left to dark and right to light. The same random orders that we used for the simultaneous group were employed. Four trials a day were given and the criterion was four successive errorless trials, i.e., 16 successive correct choices. Again, a more detailed description of the procedure used can be obtained from the original Weise and Bitterman article.

Results and Discussion

Figure 1 shows the course of learning for both groups. Following the procedure used by Weise and Bitterman we terminated each *S* as it reached criterion, and in plotting the curve we assumed that it continued to make perfect runs. Our curves show considerable resemblance to those obtained by Weise and Bitterman. They ran two analyses, one of

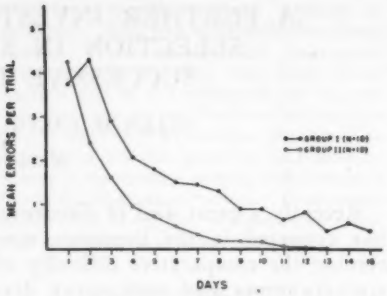

Fig. 1. The course of learning in successive (Group II) and simultaneous (Group I) discrimination

forward-going errors and one of total errors. However, since they did not give any operational definition of precisely what they considered a "retracing error," we confined our statistical analysis to forward-going errors only. Because of the difficulty of satisfying the *t*-test assumptions of trait normality and homogeneity of variance, we used Festinger's (3) nonparametric method of testing the difference in performance between Groups I and II. The difference was significant well beyond the 1% level. These results fail to confirm our hypothesis and the hypothesis suggested by Bitterman and Wodinsky.[1]

Spence has pointed out: "Just why Weise and Bitterman got opposite results is not clear, as it is difficult to interpret the very complex type of discrimination set-up they employed. The simple discrimination situation is sufficiently difficult to deal with theoretically without adding all of the problems that arise as the result of the serial nature of the multiple discrimination set-up along with the fact that it involves a gradient of reinforcement," (5, p. 91). However, Bitterman and Wodinsky (1) do not

[1] Since *S*s did not actually run or jump on the lights, it is conceivable that Bitterman and Wodinsky might consider this an invalidating factor; however, it appears doubtful that this objection would be raised.

TABLE 1

ANALYSIS OF ERRORS BY MAZE UNIT

Group	Unit 1		Unit 2		Unit 3		Unit 4	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Mean Percentage of Errors Made in Each Unit								
I	26.7	5.7	28.4	4.9	23.8	3.8	21.1	4.8
II	27.8	6.5	22.3	6.7	28.5	4.5	21.4	3.8
Mean Percentage of Errors Made in the Same Direction as the Last Rewarded Unit*								
I	71.6	10.1	63.6	8.9	65.8	14.2	69.0	16.3
II	86.3	12.1	55.0	15.5	78.3	15.0	47.4	15.0
Mean Percentage of Errors Made in the Same Direction as that of the Last Entry to the Goal Box**								
I	71.6	10.1	58.8	22.0	65.3	18.6	56.2	23.7
II	86.3	12.1	64.5	25.3	43.5	22.0	32.9	25.2

* This analysis does not include errors made in the first unit on the first trial of each day.
** This analysis does not include errors made on the first trial of each day.

believe that this is a significant factor. Even so, we decided to analyze our data to see if some type of "goal gradient" or other effect was operating differentially in the simultaneous and successive discriminations.

Table 1 presents the results of our analysis.[2] A glance at the table indicates that *S*s in Group I show a fairly consistent goal-gradient effect. They make 55% of their errors in the first two units. The difference between the percentage of errors made in the first two units and the percentage of errors made in the last two units was significant at the 1% level using Wilcoxon's nonparametric technique for paired replicates (7, p. 5). Group II does not show the same trend that Group I exhibits; these *S*s make almost the identical percentage of errors in the first two units as they do in the last two. This suggests that not only did the two groups learn at different rates, but also that their errors were made in different units.

We next examined the data to see if any pattern involving spontaneous alternation would manifest itself. Both groups show a marked tendency toward a "reverse spontaneous alternation" pattern; i.e., if an *S* was rewarded for making a left turn, he tended to turn left at the next choice point. Although both groups show this reverse spontaneous alternation pattern, the two groups again do not present parallel trends. As shown in the table, Group I has a fairly consistent pattern of alternation reversal, while Group II has a large reverse alternation effect in Units 1 and 3, but virtually no effect in Units 2 and 4. The difference between Units 1 and 3 vs. 2 and 4 is significant at the 1% level using Wilcoxon's paired replicate technique mentioned previously. It is interesting to note that Units 1 and 3 are the ones in which Group II made the most errors. Thus, the possibility of some sort of interaction between the amount

[2] The multiple-discrimination apparatus not only complicates matters theoretically, it also complicates the statistical analysis considerably. With the interaction of the large number of parameters, it is extremely difficult to statistically isolate the actual factors operating. The authors are indebted to Professors Leo Katz and I. Olkin of the mathematics department of Michigan State College for their assistance in this matter.

of errors made in a particular unit and this reverse spontaneous alternation effect is suggested, and, furthermore, it appears that this interaction acts differentially on simultaneous and successive discrimination.

The next factor that we looked for was the effect of the goal box. Both groups exhibited a strong tendency to turn in the same direction as their last entrance to the goal box. However, as in the other two analyses, the two groups do not have symmetrical trends. The data presented in the table show that Group II *Ss* are strongly influenced by the goal box in Unit 1, quite strongly influenced by it in Unit 2, and then show a "negative" influence in Units 3 and 4; i.e., in Unit 3 only 43.5% of their errors were made in the same direction as their last entrance to the goal box, and in Unit 4 an even smaller percentage, 32.9%. The difference between Units 1 and 2 vs. 3 and 4 was significant beyond the 1% level using Wilcoxon's paired replicate technique. Group I does not present a similar pattern. The *Ss* in this group are always "positively" influenced by the goal box, and the trend is relatively consistent. Therefore, an appropriate hypothesis would seem to be one which postulates an interaction between unit order and the influence of the goal box during successive discrimination but not during simultaneous.

It is apparent that Spence's statement (4) concerning the complexity of the multiple-discrimination apparatus is amply substantiated by our empirical data.

Summary

The Weise and Bitterman experiment was replicated with their maze modified so that the lamps which served as cues were placed at the end of the respective maze arms. It was hypothesized that with this breaking up of the spatial unity of the cues the simultaneous discrimination would be easier than the successive, thus reversing the results obtained by Weise and Bitterman. However, our hypothesis was not confirmed, as the successive discrimination again proved easier than the simultaneous. An analysis of our data indicated that various factors such as the goal gradient, alternation, etc. were operating differentially in the two types of discrimination. These findings and their relationship to the current position of Bitterman and his associates were discussed.

References

1. Bitterman, M. E., & Wodinsky, J. Simultaneous and successive discrimination. *Psychol. Rev.*, 1953, **60**, 371–376.
2. Calvin, A. D. Configurational learning in children. *J. educ. Psychol.*, in press.
3. Festinger, L. The significance of difference between means without reference to the frequency distribution function. *Psychometrika*, 1946, **11**, 97–105.
4. Nissen, H. W. Further comment on approach-avoidance as categories of response. *Psychol. Rev.*, 1952, **59**, 161–167.
5. Spence, K. W. The nature of response in discrimination learning. *Psychol. Rev.*, 1952, **59**, 89–93.
6. Weise, P., & Bitterman, M. E. Response selection in discriminative learning. *Psychol. Rev.*, 1951, **58**, 185–195.
7. Wilcoxon, F. *Some rapid approximate statistical procedures*. Stamford, Conn.: American Cyanamid Co., 1949.

(Received March 13, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

GENERALIZATION OF EXPECTANCY CHANGES AS A FUNCTION OF THE NATURE OF REINFORCEMENT

HAROLD M. SCHRODER AND JULIAN B. ROTTER

The Ohio State University

Brush, Bush, Jenkins, John, and Whiting (1) have pointed out recently that relatively little attention has been paid to the generalization of an extinguished instrumental response. Extinction may be considered to cover a variety of conditions that produce a decrease in the potential of a given response to occur in a given situation. Many studies have concerned themselves with a set of conditions that produce a decreased probability that a response will occur, as measured in the presence of the training cues, e.g., removing the reward in a T maze or taking a postmeasure of an attitude in the classroom after the presentation of a movie in the classroom. The purpose of this study is to inquire about the generalization of this measured change to other situations.

It is well known that under some conditions a change in the potential to respond in Situation 1 (S_1) may have little or no effect on the potential of the same response to occur in the presence of another set of cues (S_2). Such a state of affairs places limits on the overgeneralized use of such descriptions as extinction, change, inhibition, etc. This general problem is related to the potential of an organism to respond in a particular way to a given range of environmental cues and has been referred to as attitude, trait, disposition, etc.

To assist in the formulation of examples let us suppose that an expectancy (E) that a given behavior will be followed by a given reward has been learned so that when a specific cue or set of cues (S_1) is present, a given behavior (B) is elicited and directed toward that reward having a given value (RV) (**8, 9**).[1] If we ignore certain complexities, such a basic psychological event may be conceived in terms of an animal's turning left in a black T maze or described as a child's using a dependent behavior in the presence of his mother (S_1). Further, let it be assumed that this organism has learned expectancies that Reward 1 will occur in the presence of another set of cues (S_2), e.g., white T maze or teacher, or that Reward 2 will occur in the presence of another set of cues (S_2). If now a change (decrement) is somehow brought about in the expectancy for Reward 1 in S_1, the question in this paper is to predict the degree to which this measured change will generalize to S_2. Following Miller (**7**) the question may be phrased as the prediction of the slope of the generalization gradient from S_1 to S_n.

Social learning theory (**8, 9**) postulates that cues which elicit certain behaviors become functionally related if these behaviors have historically led to the same rewards. That is, if a given behavior in S_1 and another behavior in S_2 have led to the same goal, any decrease in the expectancy that

[1] This formulation bears some similarity to Tolman's sign-significate approach except that the nature and value of the reinforcement play a more important role in this social learning theory. It may also be thought of as similar to Hull's formulation involving habit strength and drive strength where the expectancy construct is substituted for habit strength and the concept of reinforcement involves only "an empirical law of effect" rather than a drive-reduction theory.

Behavior 1 will lead to Reward 1 in S_1 will generalize considerably to S_2. But if the behavior in S_2 has been historically associated with a different reward, the change in S_1 will not have as great an effect on the occurrence of Behavior 2 in S_2 assuming other known factors are constant. The specific hypothesis investigated may be stated as follows: A change in behavior in S_1 will generalize to S_2 to the degree that S_2 contains cues which have been historically associated with the same reward as the cues in S_1.

Two points should be made in reference to this hypothesis. First, factors of stimulus and response similarity will still operate to affect the slope of the gradient. The test of the hypothesis will require that these aspects are controlled. It is still mediated stimulus generalization that is the focus, but here the emphasis is mediation on the basis of common reinforcements. Second, it would be assumed that the generalization curve would become steeper as a function of increasing differentiation of reward in S_1 and S_2.

Historically there are few studies directly relating to this hypothesis. Chance (2) used two different sets of instructions to establish that two tasks were measures of the same goal for one group of human *Ss* and measures of different goals for another group. Using a level-of-aspiration paradigm she found significantly greater generalization between the two tasks when presented as measures of the same reinforcement. Crandall (3) found indications of the specificity of reinforcement effects and a tendency for the generalization curve to follow predicted lines along a dimension of need similarity where similarity was standardized by cultural "common sense" criteria. Jessor (5) controlling both expectancy and reinforcement value developed tasks, which, while unrelated to each other in terms of physical stimulus similarity, were related in terms of the functional similarity of their potential goals. Using a level-of-aspiration paradigm he demonstrated that gradients of generalization of expectancy and minimal goal were significantly steeper when the potential goal of the task was different. Although these studies present strong confirmatory data, only the study by Chance specifically attempted to control physical stimulus similarities. In her study the reinforcements were verbal and no behavioral criterion was set to establish that the goals were in fact differentiated in the experimental situation. The present study differs from the studies above in that it employs a paradigm closer to the more molecular problem of conditioning with a behavioral response (extinction) rather than a verbal statement of reduced expectancy and, of course, differs in that animal rather than human *Ss* are employed.

The basic design consists of two groups of *Ss*, two situations, and two rewards. Group A is trained to make Response 1 in the presence of one set of cues (S_1) and to expect Reward 1 and concurrently to use Response 2 in S_2 to obtain Reward 2. Group B is trained to use Response 1 in S_1 to obtain Reward 1 and concurrently to use Response 2 in S_2 to obtain Reward 1. Group A may be called the "dissimilar" reward group and Group B the "same" reward group. It is hypothesized that any change instigated in the expectancy that Response 1 will lead to Reward 1 in S_1 will generalize to Situation 2 to a greater extent for the "similar" reinforcement group, i.e., where cues in S_2 have been historically paired with the same reward. Consequently, after training to criterion both groups will be extinguished to the same criteria in S_1, and it is predicted that Group B will then take fewer trials to reach the same criteria of extinction in S_2, and will take a significantly longer running time on the first five consecutive extinction trials in S_2.

Procedure

Subjects.—The *Ss* were 24 female hooded rats approximately 90 days old at the beginning of the experiment.

Apparatus.—Two standard T mazes were used, one completely black with a fine wire mesh floor (S_1) and the other completely white with a coarse wire mesh floor (S_2). In each maze two blocks were placed between the choice point and the goal box to prevent direct visual cues; a sliding door was placed beyond the first block. These doors could be closed after *S* made its choice but before it could see the reward in the goal box. Closing the door was the criterion for a correct or incorrect response. No *S* was allowed to retrace.

TABLE 1

TRAINING AND EXTINCTION CONDITIONS FOR EACH GROUP

Group*	Reward Condition	Simultaneous Training				Extinction	
		Stage 1	Stage 2	Stage 3	Stage 4	1st	2nd
A_1	Different	LBF RWW	BF WW	LBF RWW	LBF RWW	RWW	LBF
A_2	Different	LBW RWF	BW WF	LBW RWF	LBW RWF	RWF	LBW
B_1	Same	LBF RWF	BF WF	LBF RWF	LBF RWF	RWF	LBF
B_2	Same	LBW RWW	BW WW	LBW RWW	LBW RWW	RWW	LBW

Note:—L: Left-turning response; R: right-turning response; B: black maze; W: white maze; F: food; W: water.

* $N = 6$ in each group.

Rewards.—Reward 1 was a dry powdered form of Purina dog chow. Reward 2 was a water solution with .125 gm. of saccharin added per 100 cc. of water.

Deprivation.—The *S*s were simultaneously hungry and thirsty for 20 hr. throughout all training stages and extinguishing procedures.

Groups.—Group A was composed of 12 *S*s and had different rewards counterbalanced. In Subgroup A_1 ($N = 6$) *S*s were simultaneously trained to go left in a black maze for food (LBF) and right in a white maze for water (RWW). In Subgroup A_2 ($N = 6$) *S*s were simultaneously trained to go left in a black maze for water (LBW) and right in a white maze for food (RWF).

Group B was composed of 12 *S*s and had the same rewards in the black and white mazes. In Subgroup B_1 ($N = 6$) *S*s were simultaneously trained to go left in a black maze for food (LBF) and right in a white maze for food (RWF). In Subgroup B_2 ($N = 6$) *S*s were simultaneously trained to go left in a black maze for water (LBW) and right in a white maze for water (RWW).

Training.—Since the hypothesis assumes that the rewards are historically differentiated, *S*s were given a long training procedure in finding food in one place and water in another for Group A and food in one place and food in another for Group B. Stage 1 in training (see Table 1) consisted of six sessions of 3 hr. each. The *S*s were merely placed in a starting box which led to both T mazes. The relevant rewards for each group were available in the correct arm of each T maze. For all training stages the cue-response-reward relationships for each *S* were in accordance with the procedure outlined above. In Stage 2 each *S* was given four trials a day for 12 days in respective black and white straightways. An *S* was placed in the black or white straightway following the random procedure given by Gellerman (4). This order was used throughout all trial sequences. In Stage 3 *S*s were run for five trials a day for 16 days in an L maze. At this time *S*s were being forced to make the appropriate response in each maze where the relevant cues were present. At the beginning of this training all *S*s would eat or drink for five consecutive trials when both hungry and thirsty. The final stage consisted of running *S*s in this same L maze but now with the door in the "incorrect" arm open, permitting "wrong" responses. In all cases *S*s were allowed to eat for 25 sec. and drink for 5 sec. If an incorrect choice was made, the door was closed and *S* remained in the empty goal box for 20 sec. A minimum of 2 min. was allowed between trials during training and extinction.

Extinction.—As soon as *S* reached the criterion of making the appropriate response in each maze, the extinction procedure was started. The initial extinction was carried out in the white maze for all *S*s. During the extinction the food and water containers were removed and the left-going response (incorrect for all *S*s during training) was blocked midway between the choice point and the goal box. This was to insure that *S*s only entered the empty goal box on the right during the initial extinction. All *S*s received five trials a day until they reached a criterion of 90-sec. latency. At this point they were transferred to the black maze where the trials to criterion of 90 sec. represented the measure of generalization of extinction.

For the second extinction in the black maze the right alley was blocked. The food and water containers were removed and each *S* was given five trials immediately on reaching criterion in the white maze and five trials a day until a criterion of 90-sec. latency was reached.

RESULTS

Training.—With the *F* distribution as a test of homogeneity of variance,

TABLE 2

ERRORS AND TRIALS TAKEN TO REACH CRITERION OF 18 CORRECT TRIALS OUT OF 20 IN SIMULTANEOUS TRAINING

Group*	Reward Condition	Measure	Mean Score	Sum Squared Deviations	F
A	Different	Trials	122.91	16022.90	
B	Same	Trials	117.50	5435.00	2.95
A	Different	Errors	43.58	3550.92	
B	Same	Errors	43.16	2003.47	1.85

* $N = 12$ in each group.

it was found that Subgroups A_1 and A_2 as well as Subgroups B_1 and B_2 could be drawn from a population having a common variance in reference to errors or trials to reach criterion. The application of the *t* test likewise indicated common means. Consequently Subgroups A_1 and A_2 were combined as the different-reinforcement group and B_1 and B_2 were combined as the same-reinforcement group. It may be of interest to note that the mean numbers of errors to reach criterion for Subgroups A_1, A_2, B_1, and B_2 were 38, 49, 40, and 45, respectively. When the latencies for the last five trials before reaching criterion were compared, the same-reward groups were running more slowly than the different-reward groups, but not significantly so. At Trial 40, three *S*s were discarded because of persistent position habits.

The training data for errors and trials to reach a criterion of 18/20 correct trials are given in Table 2. Although there is a tendency for Group A (different rewards) to show larger deviations in learning, these deviations do not differ significantly from those found for Group B. With 11 and 11 *df* the *F* ratio would need to be larger than 3.85 to be significant at the .05 level. These data also show that under conditions of simultaneous hunger and thirst deprivation the learning rate was similar for both groups when described in terms of mean number of errors or trials to criterion.

Extinction in Maze 1 (white).—The first two entries in Table 3 present the relevant data for this procedure. In terms of the trials taken to reach a 90-sec. latency criterion the groups may be considered as having a common variance and mean, since we would need an *F* as large as 3.85 and a *t* as large as 2.07 for the difference to be beyond that expected by chance alone. There was a tendency for the same-reward *S*s showing the longest latencies at the end of training to take more trials to reach the 90-sec. latency criterion in the extinction in Maze 1.

Extinction in Maze 2 (black).—

TABLE 3

TRIALS TO CRITERION FOR EXTINCTION IN MAZES 1 AND 2

Group*	Reward Condition	Maze	Score Unit	Mean Trials to Criterion	Sum Squared Deviations	F	t
A	Different	1	Raw	10.83	253.67		
B	Same	1	Raw	13.42	598.92	2.36	1.18
A	Different	2	Raw	12.42	180.92		
B	Same	2	Raw	5.00	36.00	5.01	5.01
A	Different	2	Transformed	3.32	4.51		
B	Same	2	Transformed	2.20	1.69	1.53	5.15

* $N = 12$ in each group.

TABLE 4

TOTAL LATENCY FOR FIRST FIVE CONSECUTIVE EXTINCTION TRIALS IN MAZE 2

Group*	Reward Condition	Score	Mean Total Time	Sum Squared Deviations	F	t
A	Different	Raw	96.91	16466.92		
B	Same	Raw	226.50	73323.00	4.45	4.97
A	Different	Transformed	9.68	37.89		
B	Same	Transformed	14.82	82.22	2.17	5.39

* $N = 12$ in each group.

These results appear in the last four entries in Table 3. On trials taken to reach a 90-sec. latency criterion in Maze 2, Group A (different-reward group) continued more than twice as long as Group B (same-reward group). The obtained t of 5.01 is significant well beyond the .01 level of confidence. However, the hypothesis of a common variance must be rejected since an F as large as 4.67 would represent the .01 level of significance as compared to an obtained F of 5.01.

Since no predictions were made in regard to this finding, any discussion would be purely post hoc. However, it is of interest to note that in Group A there is a tendency for some *S*s to extinguish very quickly in Maze 1 and take correspondingly more trials to extinguish in Maze 2. In this case, however, it would appear that the variance is correlated with the size of the mean. The last two entries in Table 2 present the data after a square root transformation. When this unit of measurement is used, the hypothesis of a common variance may be accepted. The groups differ beyond the .01 level of significance in terms of this mean score to criterion.

It may be argued that the obtained difference between means is exaggerated because five trials were given each day. If *S* failed to reach the latency criterion on the first five trials, as 11 *S*s in Group A did, some recovery would be expected 24 hr. later when the next five trials were given. Consequently all *S*s were run at least five consecutive trials in Maze 2 immediately after reaching the 90-sec. latency criterion in Maze 1. The mean or total time taken for each *S* to move from start to goal box may be used as a measure of generalization. For this analysis the latencies for the first five trials in Maze 2 were totaled. For a simple time measure an F of 4.45 was obtained which is significant at the .02 level. After a square root transformation the variance difference no longer approaches significance but the means differ at well beyond the .01 level of significance.

DISCUSSION

These results appear to support the hypothesis that a given change brought about in the potential of a response to occur in a given situation generalizes to a maximum degree to other situations, $S_2 \ldots_n$, which contain cues that have been historically paired to the same reward as that previously expected in S_1.[2]

[2] This design does not rule out an interpretation based on generalization of a specific response. It could be argued that the superior resistance to extinction in the different reward groups was due to the fact that the extinction of the left-turning response generalized more strongly from the black to the white maze in the different-reward group so that the left-turning habit interfered less with going to the right. Such an explanation seems less plausible since (*a*) the interfering habit must be weak as a result of the initial discrimination training

In this experiment factors of stimulus and response similarity were controlled apart from the goal responses of eating and drinking. The control of these factors is extremely difficult, and no attempt can be made to argue the objective merits of an expectancy versus S–R theory on the basis of these results. In conjunction with earlier experiments, the findings stress the importance of identifying the reinforcements occurring in any given situation when one is interested in any predictive problem concerning the slope of a "generalization of change" gradient. In research concerned with postmeasures in any developmental or retraining procedure, some specificity of prediction could be provided by hypothesizing in advance the situations in which one would expect maximal changes and, equally as important, those situations in which one would expect little change. Such specificity could be added at varying levels of generality by taking into account the situational cues (objects, persons, etc.), that have been historically associated with a given goal-directed behavior. An analysis of this kind may have relevance to the specificity of statements that could be made after some sampling procedures such as psychological testing.

Summary

It was hypothesized that a change in behavior in S_1 will generalize more to S_2 the more S_2 contains cues that have been historically associated with the same reinforcement as the cues in S_1. Twenty-four rats that were both hungry and thirsty throughout the experiment were divided into two major groups and each learned a different response in each of two situations simultaneously. In both groups the animals learned to turn right in a white T maze and left in a black T maze. For Group A the responses in the two situations were followed by different rewards (water and food), while for Group B the same responses were followed by the same rewards. After training to a criterion of 18/20 correct choices, one response in one maze was practically extinguished to a criterion of 90-sec. latency. Then the other response was extinguished to the same criterion in the second T maze. It was hypothesized that the effect of the extinction in Maze 1 would generalize more to Maze 2 for the group which received the same rewards, i.e., for the group for which the cues in Maze 2 had been historically paired with the same reward as in Maze 1.

The results in this limited laboratory situation offer statistical support for the hypothesis. Such findings stress the importance of identifying the nature of the reinforcement in a given situation when one is dealing with the prediction of generalized effects of change.

and (*b*) this explanation involves the unlikely hypothesis that there is more generalization to the *different* situation.

References

1. Brush, F. R., Bush, R. R., Jenkins, W. D., John, W. F., & Whiting, J. W. M. Stimulus generalization after extinction and punishment: an experimental study of displacement. *J. abnorm. soc. Psychol.*, 1952, **49**, 633–640.
2. Chance, J. E. Generalization of expectancies as a function of need relatedness. Unpublished doctor's dissertation, Ohio State Univ., 1952.
3. Crandall, V. J. A preliminary investigation of the generalization of experimentally induced frustration in fantasy production. Unpublished doctor's dissertation, Ohio State Univ., 1950.
4. Gellerman, L. W. Chance orders of alternating stimuli in visual discrimination experiments. *J. genet. Psychol.*, 1933, **42**, 207–208.
5. Jessor, R. A methodological investigation of the strength and generalization of verbal reinforcement. Unpublished doctor's dissertation, Ohio State Univ., 1951.
6. Kendler, H. H. The influence of simultaneous hunger and thirst drives upon the learning of two opposed spatial responses of the white rat. *J. exp. Psychol.*, 1946, **36**, 212–220.
7. Miller, N. E. Theory and experiment relating psychoanalytic displacement to stimulus–response generalization. *J. abnorm. soc. Psychol.*, 1948, **43**, 155–178.
8. Rotter, J. B. *Social learning and clinical psychology*. New York: Prentice Hall, 1954.
9. Rotter, J. B., Fitzgerald, B. J., & Joyce, J. A comparison of some objective measures of expectancy. *J. abnorm. soc. Psychol.*, 1954, **49**, 111–114.
10. Skinner, B. F. Are theories of learning necessary? *Psychol. Rev.*, 1950, **47**, 193–216.

(Received March 15, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

THE ROLE OF DRIVE REDUCTION IN THE CLASSICAL CONDITIONING OF AN AUTONOMICALLY MEDIATED RESPONSE[1]

DAVID ZEAMAN AND NORMA WEGNER
University of Connecticut

We propose to show that drive reduction can determine the form of a classically conditioned autonomic response, and that as a result there is no fixed relationship of anxiety and heart rate.

A number of investigators have found that the heart-rate changes during anxious anticipation of an aversive stimulus, but the direction of change has not always been the same. Some have found the heart to accelerate during anxiety (**1, 2, 5, 13, 17**), while others have recorded a slowdown (**5, 11, 13, 16, 19**). A resolution of these discrepant results is offered in the present study.

The most extensive work on heart conditioning has been done recently by Notterman, Schoenfeld, and Bersh (**10, 11, 12**). They have shown that when anxiety is generated by a tone-shock conditioning procedure the heart slows in rate during the anxiety period between tone and shock. Zeaman, Deane, and Wegner (**19**) replicated this finding and made a detailed beat-by-beat analysis of the form of both the CR and UR. It is reported in the latter study that during the 6-sec. CS-US interval, the heart-rate CR to the 1-sec. tone took the form of a slight initial rise followed by a deceleration of greater magnitude and duration. In contrast, the UR to the 6-sec. shock was a sharp elevation in rate reaching a peak in 3 sec., and then a progressive drop in rate until the point of shock termination. The over-all picture was that of a slowdown CR and speed-up UR. An attempt was made to relate the findings to several theoretical formulations of the mechanism of conditioning. One of these, the drive-reduction view, suggested that the predominantly decelerative form of the CR was due to the fact that the heart was decelerating at the time of maximal drive reduction (shock-off). A direct test of this view was proposed. It consisted of shortening the duration of the shock US to 2 sec., so that drive reduction would now occur during the acceleration stage of the UR. The predicted outcome was a change in the form of the CR from deceleration to acceleration. The present experiment reports the result.

METHOD

Apparatus

The apparatus has been described in detail in a previous report (**19**). Briefly, it consisted of an electrocardiograph to measure the heart rate, a 13 V.A.C. shock for the US, a tone (60 db, 512 c.p.s.) for the CS, and electronic timers for controlling the sequence of stimuli.

Subjects

The *S*s were 16 students at the University of Connecticut, 10 males and 6 females. All were paid volunteers.

Procedure

The classical trace-conditioning procedure employed was very similar to that described in an earlier paper (**19**); the major change was the shortening of the duration of the shock from 6 sec. to 2 sec. Our technique was closely patterned after that of Notterman, Schoenfeld, and Bersh (**11**).

[1] This is Technical Report No. 8 under Contract Nonr-631 (00) between the Office of Naval Research and the University of Connecticut.

Preconditioning.—All *Ss* were given a series of preconditioning trials consisting of the presentation of a number of 1-sec. tones spaced at irregular intervals of 1 to 2 min. Ten of our *Ss* were given 10 such preconditioning trials, and six were given 20, before the start of the conditioning trials.

Conditioning.—A conditioning trial consisted of a 1-sec. tone followed in 6 sec. by a 2-sec. shock. All *Ss* were given 11 conditioning trials separated by intertrial intervals which varied irregularly between 1 and 2 min.

Cardiograph records were taken on each trial from about 20 sec. before the tone to about 40 sec. after the tone. The *Ss* were instructed that they would receive a series of tones and a series of shocks but the order was not mentioned. The instructions were identical to those of Notterman, Schoenfeld, and Bersh (11).

Results

Preconditioning

The response of the heart to tone alone in the preconditioning trials is represented by the dashed line in the lower half of Fig. 1.[2] The response, an unconditioned one, is an immediate drop in rate of about 2 b/m, with partial recovery during the 10 posttone beats. The curve is an average of all *Ss* for all preconditioning trials. Since there is no appreciable adaptation of this response, the average curve of the group receiving only 10 preconditioning trials differed only slightly from that of the group receiving 20. For this reason, the results of both groups have been pooled.

Conditioning

The effect of tone-shock pairing on the heart's response to tone is shown by the solid curve in the lower half of Fig. 1. The major effect here, after a brief tone-UR dip, is a rise of about 4 b/m followed by a fall to a level below that of the pretone period. The last two posttone beats are the next-to-last and last beats before shock. They have an average ordinal rank of 6.8 and 7.8 counting in the direction of tone to shock. The shock brings about a sharp elevation in rate of about 6½ b/m. At the time of shock-off the heart is strongly accelerating. All points on the solid-line curve are the averages of all 16 *Ss* for Conditioning Trials 2–11.

A corrected measure of the effects of conditioning alone is presented in the top half of Fig. 1; it is the difference between the conditioning and preconditioning curves in the lower half of Fig. 1. This correction is based on the assumption that the tone UR, which remained constant during the preconditioning trials, continues to have a constant effect on heart rate during the conditioning trials. Thus, the tone UR can be subtracted out of the conditioning curve which presumably contains both the tone UR and CR. The result of this correction is to bring out the predominantly accelerative form of

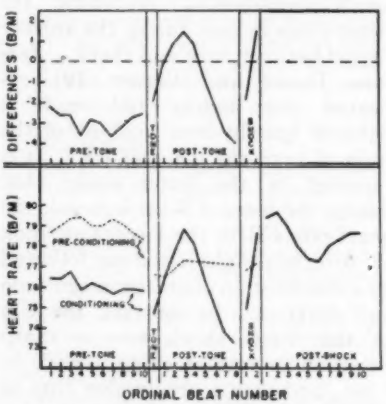

FIG. 1. The form of the heart-rate response during preconditioning and conditioning (lower half). The top half represents the difference between the two lower curves.

[2] Tables of data which form the basis for these figures have been deposited with ADI. Order document No. 4402; $1.25 for 35 mm. microfilm and $1.25 for photocopies.

the CR. Six of our 16 *Ss* had *uncorrected* CR's resembling closely the average *corrected* CR, that is, with no drop below pretone level for the next-to-last and last beats before shock. The negative values of the corrected curve signify only that the over-all level of heart rate is lower during the conditioning trials than the preconditioning trials, because of *Ss'* progressive approach to a basal state during the course of the experimental session.

Having shown that the form of the CR is an acceleration with a maximum amplitude at the third posttone beat, we show now that this CR has a gradual onset. The course of conditioning can be traced in Fig. 2. On each trial, the mean rates, for all *Ss*, of beats 3 and 4 are averaged and expressed as a difference from the mean rate during the 10 pretone beats on that trial. The difference measure was chosen here merely to cancel out trial-by-trial changes in over-all level of heart rate. The solid line shows a typical negatively accelerated curve of conditioning. The negative values of the difference measure during the ten preconditioning trials indicate that the tone UR has depressed the rate below pretone level by approximately 1½ b/m. It can be seen from the flatness of the preconditioning curve that the tone UR showed on tendency to adapt out. For the subgroup which received 20 preconditioning trials, only the data from the second 10 of the preconditioning trials are included in Fig. 2.

The next-to-last and last beats before shock were also averaged and expressed as a difference from the

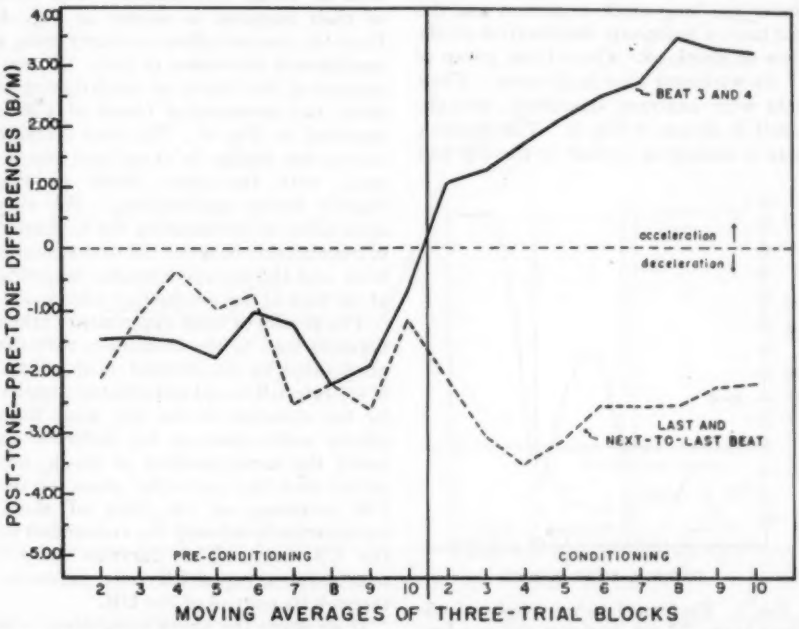

FIG. 2. The course of conditioning for the accelerated and decelerated portions of the CR

mean of the 10 pretone beats on each trial. Considering the amount of preconditioning variability, we see no convincing evidence of a conditioned deceleration for this portion of the response.

Discussion

The data we have presented clearly establish the fact that the heart shows a conditioned acceleration in anticipation of a short shock. This is to be contrasted with the previous finding of a conditioned deceleration when a long shock was used (11, 19). What the heart is doing at the time of shock-off seems to make the difference.

In the light of this finding, we were led to re-examine the data from our long-shock experiment, because we had observed there some evidence of a slight conditioned acceleration in the beginning of the CR, just before the large conditioned deceleration. The possibility was considered that some *S*s were atypical under the long-shock condition and did not have a heart-rate deceleration at the time of shock-off. Out of the group of 17 *S*s we found four such cases. Their data were analyzed separately, and the result is shown in Fig. 3. The measure here is analogous to that in the top half

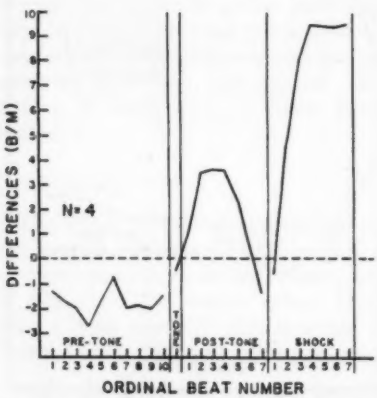

Fig. 3. The form of the heart response for *S*s whose hearts did not decelerate during a long shock

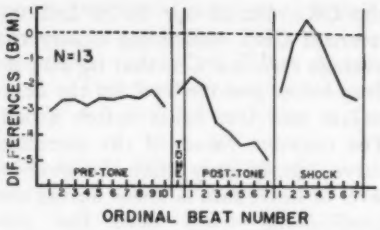

Fig. 4. The form of the heart response for *S*s whose hearts did decelerate during a long shock

of Fig. 1. Note that for these *S*s the heart, at the time of shock-off, has just passed its peak of acceleration, and has not yet started its deceleration. Observe that, more importantly, the CR resembles closely that found in the present experiment with short shock. It must be emphasized that these *S*s were selected only on the basis of the form of UR, not CR.

The remaining 13 *S*s of our long-shock experiment were all decelerating at the time of shock termination. The form of their response is shown in Fig. 4. Here the reverse effect is clearly seen, a conditioned slowdown in rate. A comparison of the course of conditioning of these two contrasting forms of CR is depicted in Fig. 5. The two learning curves are similar in shape and magnitude, with the short shock yielding slightly better conditioning. We view each curve as representing the formation of a connection between the tone stimulus trace and the specific response occurring at the time of drive reduction (shock-off).

The results of both experiments taken together lead to the conclusion that the accelerated or decelerated form of the heart-rate CR is not determined directly by the duration of the US, since both effects were obtained for different *S*s under the same duration of shock, but rather that the particular phase of the UR occurring at the time of shock termination is selected for connection to the CS trace. The duration of US selects the type of CR only indirectly through its control of the UR.

If we grant the above conclusion, it is not difficult to understand how previous

investigators found sometimes one and sometimes another type of heart-rate CR. A variety of types and durations of unconditioned stimuli has been used with no attention to the important factor of drive reduction, which predicts the type of CR to expect. Earlier studies in cardiac conditioning have not been reported in sufficient detail to allow us to use them as precise tests of our conclusion. No obviously contrary cases have been reported.

We wish to consider briefly the relation of our results to theories of anxiety and the mechanism of conditioning. If one holds to the simple descriptive definition of anxiety as a state produced by some current stimulus which in the past has been followed by a noxious stimulus (4, 15), then the different heart-rate CR's in both our short-shock and long-shock experiments can be said to have occurred during anxiety. Thus the appropriate statement of the relationship of anxiety and heart rate is that the heart changes its rate during anxiety, but in a direction determined by factors other than anxiety.

If one views anxiety as an autonomically mediated response, conditioned like other viscero-vascular responses, then the mechanism by which heart rate is conditioned will also be the mechanism by which anxiety is learned. There are at least two schools of thought on this matter: one which holds that autonomic responses are conditioned according to a contiguity principle (3, 7, 9, 14, 18), and another which maintains that they are conditioned by drive reduction (6, 8). Since in the present experiment it was the shock termination (drive reduction) which proved to be the critical element in selecting the autonomic response to be conditioned, the results lend most comfort to the drive-reduction view.

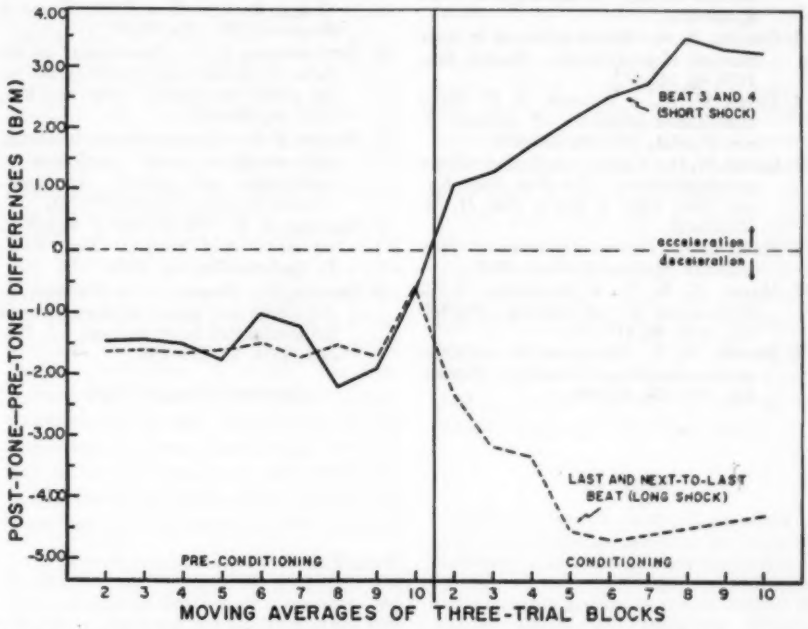

FIG. 5. The acquisition of the conditioned acceleration in the short-shock experiment compared with that of the conditoned deceleration in the long-shock experiment

Summary

The heart-rate response of human *Ss* was classically conditioned, using a 1-sec. tone CS and a 2-sec. shock US separated by a 6-sec. delay interval. During the anxiety period between tone and shock, the conditioned cardiac response appeared as an acceleration in rate. This was contrasted with the conditioned deceleration found during anxiety when a long (6-sec.) shock was used. The difference in the form of the two CR's was attributed to the fact that at the time of shock termination the heart was accelerating in the short-shock condition and decelerating in the long-shock condition. The results are seen to favor the view that the mechanism of conditioning for this autonomically mediated response was that of drive reduction.

References

1. Anderson, O. D., & Parmenter, R. A long-term study of the experimental neurosis in the sheep and dog. *Psychosom. Med. Monogr.*, 1941, **2**, Nos. 3 & 4.
2. Blatz, W. E. The cardiac, respiratory, and electrical phenomena involved in the emotion of fear. *J. exp. Psychol.*, 1925, **8**, 109–132.
3. Culler, E. A. Recent advances in some concepts of conditioning. *Psychol. Rev.*, 1938, **45**, 143–153.
4. Estes, W. K., & Skinner, B. F. Some quantitative properties of anxiety. *J. exp. Psychol.*, 1941, **29**, 390–400.
5. Gantt, W. H. Cardiac conditioned reflexes to painful stimuli. *Fed. Proc. Amer. Soc. exp. Biol.*, 1942, **1**, No. 1, Part II, 28. (Abstract)
6. Hull, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
7. Maier, N. R. F., & Schneirla, T. C. Mechanisms in conditioning. *Psychol. Rev.*, 1942, **49**, 117–134.
8. Miller, N. E. Comments on multiple-process conceptions of learning. *Psychol. Rev.*, 1951, **58**, 375–381.
9. Mowrer, O. H. *Learning theory and personality dynamics*. New York: Ronald, 1950.
10. Notterman, J. M., Schoenfeld, W. N., & Bersh, P. J. A comparison of three extinction procedures following heart rate conditioning. *J. abnorm. soc. Psychol.*, 1952, **47**, 674–677.
11. Notterman, J. M., Schoenfeld, W. N., & Bersh, P. J. Conditioned heart rate response in human beings during experimental anxiety. *J. comp. physiol. Psychol.*, 1952, **45**, 1–8.
12. Notterman, J. M., Schoenfeld, W. N., & Bersh, P. J. Partial reinforcement and conditioned heart rate response in human subjects. *Science*, 1952, **115**, 77–79
13. Robinson, J., & Gantt, W. H. The cardiac component of the orienting reflex. *Fed. Proc. Amer. Soc. exp. Biol.*, 1946, **5**, No. 1, Part 2, 87–88. (Abstract)
14. Schlosberg, H. The relationship between success and the laws of conditioning. *Psychol. Rev.*, 1937, **44**, 379–394.
15. Schoenfeld, W. N. An experimental approach to anxiety, escape, and avoidance behavior. In P. H. Hoch & J. Zubin (Eds.), *Anxiety*. New York: Grune & Stratton, 1950. Pp. 70–99.
16. Sherrington, C. S. Experiments on the value of vascular and visceral factors for the genesis of emotion. *Proc. roy. Soc.*, 1900, **66**, 390–403.
17. Skaggs, E. B. Changes in pulse, breathing, and steadiness under conditions of startledness and excited expectancy. *J. comp. Psychol.*, 1926, **6**, 303–318.
18. Skinner, B. F. *The behavior of organisms: an experimental analysis*. New York: D. Appleton-Century, 1938.
19. Zeaman, D., Deane, G., & Wegner, N. Amplitude and latency characteristics of the conditioned heart response. *J. Psychol.*, 1954, **38**, 235–250.

(Received March 18, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

INFLUENCE OF AWARENESS OF REINFORCEMENT ON VERBAL CONDITIONING[1]

JOSEPH B. SIDOWSKI

University of Wisconsin

According to the law of effect, learning may take place in a relatively automatic way (11). If we use *S*'s ability to report as a criterion of awareness, we may say that an *S* may learn without awareness of the responses he learns, or of the motivation or reinforcement involved.

Early studies by Thorndike (14, 15) presented evidence that a satisfying aftereffect could strengthen the connection which it followed in cases where *S* was unaware of what the connection was. The validity of this evidence rested upon the assumption that lack of awareness was indicated by continuous slow learning. Criticizing the criterion used by Thorndike, other experimenters (7, 8, 9) presented evidence which left unverified the assumption of lack of awareness which was essential for Thorndike. Additional results (12) showed only little evidence for learning without awareness.

In a more recent study investigating the effects of a reinforcing stimulus upon verbal behavior, Greenspoon (6) concluded that presenting a stimulus such as "umm-hmm" after plural verbal responses increased the frequency of plural responses. Upon questioning at the conclusion of the experiment, it was found that most of the *Ss* were unaware of this increase.

Following a procedure similar to that used by Greenspoon, the present experiment was designed to study the influence of a stimulus light blink following plural responses upon the rate of plural responses. The *Ss* in all groups were merely told to say words. The plural responses given by *Ss* in the experimental groups were followed by a light blink in an attempt to condition this verbal response class. Our first hypothesis was that the introduction of a stimulus light following a plural response would increase the probability of that response occurring, and that this result would be obtained even for *Ss* who were not aware of the purpose of the experiment. The *Ss*' ability to report the purpose of the experiment, including a statement concerning the class of words being conditioned (plurals), was used as the criterion of awareness. An additional group, the Task group, was included to determine whether *Ss* who were given a problem to solve would condition at a faster rate. These *Ss* were told that the light could be made to blink, whereas *Ss* in the other groups were not even told that the light was part of the experiment. As a second hypothesis, it was expected that the Task group would give a larger number of plural responses than the Nontask group.

METHOD

Subjects.—The *Ss* were 62 members of elementary psychology courses at the University of Wisconsin, and they received class points for serving. Forty-seven *Ss* were drawn from the spring semester population and 15 *Ss* from the summer session. No significant difference ($F = .079$) was found between the two population samples. Therefore, the data from both

[1] The writer is indebted to Drs. L. B. Wyckoff, R. Peterson, C. Nystrom, E. J. Archer, and J. Greenspoon for their helpful suggestions. An abstract of this paper was presented at the 1954 meeting of the Midwestern Psychological Association.

populations were pooled in the subsequent analysis.

Six *Ss* were omitted from the main analysis because they failed to meet certain criteria which had been established a priori. These *Ss* will be discussed later. The remaining 56 *Ss* were randomly divided among the four groups with 19 *Ss* in the Nontask, 13 *Ss* in the Task, 15 *Ss* in the Nonreinforcement, and 9 *Ss* in the Random Blink groups.

Procedure.—Each *S* was seated in one corner of the experimental room facing the wall and a 7-w. light bulb mounted at approximately eye height. The *E* sat at a table in back of *S* where he recorded plural and nonplural responses on mimeographed sheets and manipulated a foot pedal which controlled the blinking of the light. At no time during the experiment was *S* permitted to turn around to face *E*.

The *Ss* were randomly assigned to four groups. The Nontask group was given the following instructions: "When I tell you to begin, I want you to start saying words. Any words will do but you are not to count, say sentences, or phrases. Remember, that you are to keep saying words until you are told to stop. All right, go ahead." The same instructions were also used for the two control groups described below. The Task group received the same instructions with the addition that each *S* was told to notice the stimulus light bulb which could be made to blink, and that it was his job to make the light blink as many times as possible during the course of the experiment. Each plural response spoken by any *S* in either of the two experimental groups, Nontask and Task, was followed by a blink of the light.

Two control groups were also employed. The Nonreinforcement group merely said words while *E* recorded each word spoken as either a plural or a nonplural response. At no time was the light presented for this group. A Random Blink group had a light presented randomly throughout the session regardless of the type of word response which had preceded it. The frequency of blinking was chosen so that the light blinked approximately the same number of times for this group as for the two experimental groups.

Each *S* was required to give a total of 325 verbal responses during the experimental session. The 325 total responses were divided into five continuous trials of 65 responses each. The *S* was not interrupted during the course of the experiment unless he repeated the same word three times in succession at which time *E* merely said, "Try not to repeat," which was usually sufficient. Any questions asked by *S* during the course of the experiment were ignored.

After the 325 responses were completed, each *S* was asked the following two questions: "Were you aware of the purpose of the experiment?" and, "Were you aware of the purpose of the light?" The scores of *Ss* who correctly stated the purpose of the light and the purpose of the experiment, to the extent that the purpose was to condition plural words, were omitted from the main analysis. Six *Ss* correctly stated the purpose of the experiment.

Results

The following tabulations of the experimental data are based on the results of the 56 *Ss* who failed to state the purpose of the experiment.

The essential findings of the experiment are shown in Fig. 1. In Fig. 1 the mean number of plural responses is plotted for successive trials for each of the four groups. Each trial designates a period during which each *S* gave 65 total responses, with the experimental session being divided into 5 such trials.

Results of the four groups.—Before carrying out an Alexander trend analysis (1), the data were subjected to a two-tailed *F* test for homogeneity of variance. The homogeneity test was carried out on the variances of the pooled experimental and the pooled control groups. An *F* of 3.18 was found which was significant beyond the .01 level of confidence and indicated heterogeneity of variance. Lindquist (10) suggests that even in

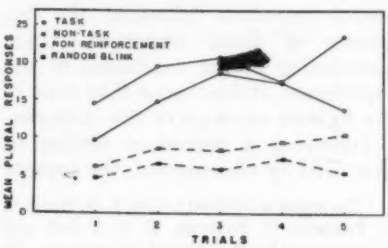

Fig. 1. Mean number of plural responses for the 56 *Ss* who were unable to report the purpose of the experiment

the presence of heterogeneity of variance it may be permissible to use the trend analysis provided a slightly higher "apparent" level of significance is used. The higher level, the 1% level of confidence rather than the 5% level, was used in analyzing the data of this experiment.

An Alexander trend analysis for the number of plural responses for the four conditions is summarized in Table 1. The *F* of 7.62 in line *a* of Table 1 indicates that the differences between the means of the four groups are significantly greater than could be accounted for by individual differences. The *F* of .33 in line *d* is nonsignificant and indicates that between-group slopes do not differ beyond what could be expected on the basis of differences between individual slopes. The significant difference in line *f*, *F* of 13.09, indicates that there was significant over-all trend.

Duncan's test for differences between ranked treatments (5) was carried out on the means of the four groups. This was done in order to determine the significance of the differences between the ranked means of each of the four groups. The results of this analysis are presented in Table 2. The Task group mean was found to be significantly higher than all other group means except the Nontask group mean. The Nontask group mean was significantly higher than either of the control group means. The difference between the Random Blink and Nonreinforcement group means was nonsignificant.

TABLE 1

SUMMARY OF ALEXANDER TREND ANALYSIS OF THE NUMBER OF PLURAL RESPONSES FOR THE FOUR GROUPS

Source of Variation	Error Term	*df*	Mean Square	*F*
a. Between-group means	*b*	3	2,095.33	7.62***
b. Between-individual means	*h*	52	275.15	6.60***
c. Group deviations from estimation	*h*	9	65.89	1.58
d. Between-group slopes	*e*	3	30.67	.33
e. Between-individual slopes	*h*	52	93.60	2.24**
f. Over-all slope	*h*	1	546.00	13.09***
g. Over-all deviations from linearity	*h*	3	134.33	3.22
h. Individual deviations from estimation		156	41.70	

** Significant beyond .01 level of confidence.
*** Significant beyond .001 level of confidence.

TABLE 2

DUNCAN'S TEST FOR DIFFERENCES BETWEEN RANKED TREATMENTS

Ranks	Means
Random Blink	30.00
Nonreinforcement	42.60
Nontask	74.84
Task	94.31

Differences between Ranked Means

Groups	Difference
Task-Random Blink	64.31***
Task-Nonreinforcement	51.71***
Task-Nontask	19.47
Nontask-Random Blink	44.84***
Nontask-Nonreinforcement	32.24***
Nonreinforcement-Random Blink	12.60

*** Significant beyond .001 level of confidence.

Results of the combined experimental and combined control groups.—The purpose of combining the above groups was to further test the hypothesis that the presentation of the light following plurals acted as a reinforcing stimulus for the Task and Nontask groups. This test also gives an indication of whether or not the experimental and control group slopes and higher order components differed (Table 3).

The *F* of 19.79 in line *a* of Table 3 indicates that the differences between the means of the combined Task and

TABLE 3

SUMMARY OF ALEXANDER TREND ANALYSIS OF THE NUMBER OF PLURAL RESPONSES FOR THE COMBINED CONTROL AND COMBINED EXPERIMENTAL GROUPS

Source of Variation	Error Term	df	Mean Square	F
a. Between-group means	b	1	5,523.00	19.79***
b. Between-individual means	h	54	279.09	6.62***
c. Group deviations from estimation	h	3	54.67	1.30
d. Between-group slopes	e	1	58.00	.64
e. Between-individual slopes	h	54	90.76	2.15**
f. Over-all slope	h	1	546.00	12.94***
g. Over-all deviations from linearity	h	3	134.33	3.18
h. Individual deviations from estimation		162	42.19	

** Significant beyond .01 level of confidence.
*** Significant beyond .001 level of confidence.

Nontask and the means of the combined Nonreinforcement and Random Blink groups were significantly greater than could be accounted for by individual differences. No significant difference was found between group slopes, line *d*. The over-all slope term, line *f*, was significant, which indicates that there was a significant linear increase in the number of plural responses made by both groups.

Results of analysis between the Task and Nontask groups.—An Alexander trend analysis (1) of the number of plural responses for the Task and Nontask groups was performed and the results are shown in Table 4. This analysis was carried out in order to test the hypothesis that *Ss* presented with a task would condition at a faster rate than those *Ss* in the Nontask group. No significant differences were found between either the group means or the group slopes.

TABLE 4

SUMMARY OF ALEXANDER TREND ANALYSIS OF THE NUMBER OF PLURAL RESPONSES FOR THE TASK AND NONTASK GROUPS

Source of Variation	Error Term	df	Mean Square	F
a. Between-group means	b	1	584.00	1.47
b. Between-individual means	h	30	397.50	7.89***
c. Group deviations from estimation	h	3	138.33	2.74
d. Between-group slopes	e	1	7.00	.05
e. Between-individual slopes	h	30	147.17	2.92**
f. Over-all slope	h	1	513.00	10.18**
g. Over-all deviations from linearity	h	3	174.00	3.45
h. Individual deviations from estimation		90	50.40	

** Significant beyond .01 level of confidence.
*** Significant beyond .001 level of confidence.

Results of six Ss who met the criterion of awareness.—Six *Ss* correctly stated the purpose of the experiment. Three *Ss* were from the Task group and three from the Nontask group. As is shown in Fig. 2, the average number of plural responses for the three *Ss* in the Task group is higher than that of any of the main experimental groups. The average curve of the three *Ss* in the Nontask group is about the same as the curves for the two main control

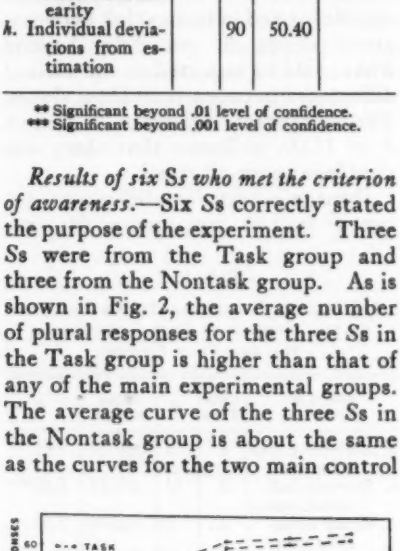

FIG. 2. Number of plural responses for the six *Ss* reporting awareness of the purpose of the experiment

groups. It is of interest to note that only one of the six *Ss* who stated the purpose of the experiment apparently gained sudden "insight" into the solution of the problem.

Discussion

The results of the present experiment demonstrate the effectiveness of the light blink as a reinforcing stimulus. This is indicated in several of the analyses: (*a*) a significant difference is shown between the four group means, (*b*) Duncan's test (5) indicates that the means of the Task and of the Nontask groups were significantly higher than either of the control group means, and (*c*) a significant difference was found between group means for the combined experimental and the combined control groups. These findings support our first hypothesis that a light blink presented following plural responses would increase the probability of plurals occurring. If we accept *S*'s ability to report the purpose of the experiment as a criterion of awareness, we may say that learning without awareness has been demonstrated since this difference between groups was established on the basis of the 56 *Ss* who failed to meet this requirement. This tends to support the law of effect in showing that learning may take place without *S* being aware of the responses he learns, or of the reinforcement involved. Implications for this type of learning for clinical psychology have been set forth by Dollard and Miller (4). It is of importance to note that although differences in the conditioning of plural responses were found between the experimental and control groups, that not all *Ss* in the experimental groups were conditioned to the light.

Our second hypothesis stated that the number of plural responses would be greater for the Task group than for the Nontask group. A difference in the expected direction was obtained but this difference was not statistically significant. This finding leaves the validity of the hypothesis in doubt. In any event, it can be said that if the task instructions had any effect on performance, the effect was small in comparison to the variability between *Ss*. Consideration of experimental findings in the area of intentional vs. incidental learning (2, 11, 13) suggests one reason why a more appreciable difference failed to appear. Results (2, 13) indicate that when learning occurs in the absence of experimentally administered instructions to learn, it is frequently the result of self-instructed "sets" brought into the situation by *S*. In the present experiment, it was noted that some of the *Ss* in the Nontask group tended to instruct themselves and to set up a problem-solving situation of their own. This was partially borne out in statements made by some of the *Ss*, such as: "I thought that you were trying to get me to say words which were associated with some of my personal problems," or "You were trying to see how many words I could say," or "You were trying to note my reactions to some of the words that I was saying." This tends to support the above findings relating to self-induced sets. The results do not, however, support the conclusions of Brown (3) who found that intentional learning was consistently superior to incidental learning. It should be noted, however, that although the statistical analysis indicates that differences between the Task and Nontask groups may be due to chance, an examination of the curves (Fig. 1) suggests that a real difference in favor of the Task group might have been obtained if the experiment had been continued for a longer period of time.

In conclusion, the findings of the present experiment confirm Greenspoon's (6) results concerning the influence of a reinforcing stimulus upon verbal behavior. They provide us with information to the effect that: (*a*) an incidental stimulus, such as a light blink, may be used to strengthen certain classes of verbal responses, with the *S* being unaware of what he is learning, and, in addition, (*b*) that the assigning of a task may have little or no effect on an *S*'s behavior. If the task assignment had any

effect, it was too small to be detected under conditions of the present experiment.

Summary

The purpose of this experiment was to study the influence of a stimulus light blink following plural responses upon the rate of plural responses. The *Ss* were instructed to say words and a light blink was presented following plural responses in the experimental groups. It was hypothesized that the conditioning of plurals would take place even for *Ss* who were not aware of the purpose of the experiment. The *Ss*' ability to report the purpose of the experiment and the purpose of the light, including a statement concerning the class of words being conditioned, was used as the criterion of awareness.

Sixty-two *Ss* were assigned to four groups. An experimental group—the Nontask group—and two control groups were not told that the light was part of the experiment. The fourth group, a Task group, was told that the light could be made to blink and that it was their job to make it blink. It was hypothesized, as a second hypothesis, that the Task group, given a problem to solve, would condition at a faster rate than the Nontask group.

The results indicated the effectiveness of the light flash as a reinforcing stimulus. Learning without awareness was demonstrated since the differences between the experimental and control groups was established on the basis of the 56 *Ss* who failed to state the purpose of the experiment. The Task group did not make a significantly larger number of plural responses than the Nontask group, which indicates that specifically assigning a problem for *S* to solve had relatively little effect on behavior, as measured in this study.

References

1. Alexander, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.
2. Bahrick, H. P., Ranken, R. E., & Fitts, P. M. Effect of incentive upon reactions to peripheral stimuli. *J. exp. Psychol.*, 1952, **44**, 400–406.
3. Brown, G. Factors influencing incidental learning. *J. exp. Psychol.*, 1954, **47**, 163–164.
4. Dollard, J., & Miller, N. E. *Personality and psychotherapy.* New York: McGraw-Hill, 1950.
5. Duncan, D. B. A significance test for differences between ranked treatments in an analysis of variance. *Virginia J. Sci.*, 1951, **2**, 171–189.
6. Greenspoon, J. The effect of verbal and non-verbal stimuli on the frequency of members of two verbal response classes. Unpublished doctor's dissertation, Univ. Indiana, 1950.
7. Herrick, C. J. Awareness of the inductive solution of problems using words as stimuli. *Arch. Psychol., N. Y.*, 1939, **33**, No. 234.
8. Irwin, F. W., Kaufman, K., Prior, G., & Weaver, H. On "learning without awareness of what is being learned." *J. exp. Psychol.*, 1934, **17**, 823–827.
9. Irwin, F. W. A rejoinder. *J. exp. Psychol.*, 1935, **18**, 389.
10. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
11. McGeoch, J. A., & Irion, A. L. *The psychology of human learning.* New York: Longmans, Green, 1952.
12. Postman, L., & Jarrett, R. An experimental analysis of learning without awareness. *Amer. J. Psychol.*, 1952, **65**, 244–255.
13. Postman, L., & Senders, V. L. Incidental learning and generality of set. *J. exp. Psychol.*, 1946, **36**, 153–165.
14. Thorndike, E. L. *The fundamentals of learning.* New York: Teachers Coll., Columbia Univ., 1932.
15. Thorndike, E. L., & Rock, R. T. Learning without awareness of what is being learned or intent to learn it. *J. exp. Psychol.*, 1934, **17**, 1–19.

(Received February 17, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

THE RELATIONSHIP UNDER STRESS BETWEEN CHANGES IN SKIN TEMPERATURE, ELECTRICAL SKIN RESISTANCE, AND PULSE RATE

LAWRENCE M. BAKER AND WILLIAM M. TAYLOR

Purdue University

The measurement of skin temperature has been attempted for a variety of purposes and through numerous techniques. Early efforts were concerned with the reactions at the extremities in response to nicotine stimulation (**7, 13, 17, 18**). A continuing interest in skin temperature measures has been centered in problems of radiation, ventilation, and factors of human metabolism (**9**). However, inference of the presence or absence of emotional states from the temperature response was essentially unsuccessful when first undertaken by Helson and Quantius (**11**). More recently a number of investigators have attempted to establish relationships between surface temperature and (*a*) emotional conflict states, broadly defined (**15, 16**); (*b*) social traits in children (**5**); (*c*) changes in affective states (**12, 14, 19**). The lack of precise and practical instruments for measurement has been an important limitation of many of these studies.

Almost all the investigations have considered skin temperature changes to be of a vasomotor origin. Increases in skin temperature are generally conceived as being the result of vasodilatation of arterioles, whereby cutaneous blood flow is increased and the temperature rises. A decrease in skin temperature is said to follow vasoconstriction. The association of cold hands with insecurity and warm hands during emotional security is a parallel here. However, the vasomotor explanation of skin temperature changes, though having the general support of the physiologists (**2, 3, 4**), is suspect not only by inconsistent results of studies cited, but by exceptions of the observations of physiologists themselves. Pale skin with constricted superficial vessels may be even hotter than red skin. This would occur when blood flow is rapid in deeper vessels and the resultant heat is conducted through overlying skin for surface dissipation via radiation (**3**). Lacking, then, is objective information as to the basis of skin temperature changes which may occur, as well as accurate instrumentation for determining such changes.

The purpose of this experiment was to study the relationships between skin temperature changes and other physiological changes previously associated with emotional responses. To this end it was necessary to devise a more accurate skin temperature recording apparatus. Electrical skin resistance and pulse rate were the other physiological changes involved.

Method

Apparatus.—A multichannel recorder, consisting of two electrical channels and one pneumatic channel, was utilized. The relatively new and recently further improved heat measuring device, the thermistor, was adapted for the temperature pickup. This unit and its arrangement for recording have been described by Baker and Taylor (**1**). The skin resistance electrodes consisted of two small, plastic, cuplike units, the bottoms of which were made of zinc. These electrodes were fastened to a hinged holder which has a spring-type action that permitted light and uniform pressure on the skin surface by each electrode. Through this

means it was possible to reduce the influence of extraneous movements as the method allowed the electrodes to be placed between the second and third joints of the middle finger while the hand was cupped and relaxed over a form-fitting board. The blood pressure cuff was of the snap-on type and was attached to the left leg just above the ankle bone. This position was chosen as preliminary checking indicated that longer recordings could be made (cuff inflated below systolic point) with apparently reduced pain or psychological disturbance of the subjects.

Stimulus.—A stress stimulation of assumed emotion-provoking strength was achieved by means of a large induction coil with two electrodes arranged so as to produce a display of jumping electrical sparks of about 3 in. in length. The apparatus was located just to the right and at near shoulder level to *S*. It was shielded from sight by a wooden frame; only the electrodes projected through a slit. The unit was seldom noticed during the experimental period and therefore the stimulus came as a distinct "surprise" or "shock" when presented. No contact or electric shock was administered to *S*. A second stress stimulation was obtained by use of a hand dynamometer. An optimal level of induced muscular tension was determined for each *S* before the beginning of the experiment by the Courts method (6). Therefore, during the actual stimulus period, *S* maintained induced tension by the hand dynamometer at one-fourth of his maximum level. As a precaution to ascertain that *S* was maintaining the required level of induced tension, an adjustable contact point was attached to the dynamometer which completed a circuit to a small bulb which in turn remained lighted if correct tension was being maintained. The bulb was always visible to *S* and *E*.

Testing room.—The walls were covered with burlap and Celotex materials which effectively reduced outside noises. It is believed that the room was sufficiently soundproof for our purposes. A large armchair with cushions served as a seating arrangement for *S*. The arms of this chair were lined with form-fitting foam rubber mats which were conducive to a relaxed, upright position. This also was intended to reduce movements of the arms and hands. A footstool was provided at the base of the chair so that the feet of *S* could be maintained in a relatively comfortable position with a minimum of movement.

Procedure.—In an outer room *S* was given an explanation of the general nature of the experiment and then shown the principal recording unit in the testing room so as to gain rapport and lessen anxiety. Following this, he was seated in the large armchair and the apparatus was attached in the following manner:

The thermistor unit was placed against the palm of the left hand, because of the rich capillary bed there. It was held in place by an adjustable elastic band which permitted good thermistor contact with the skin, but avoided undue pressure.

Next, the skin resistance electrodes were attached to the middle finger of the left hand, following the plan for a relaxed position of the hand. The electrodes were always thoroughly cleaned, and zinc sulphate paste of approximately the same moisture content was applied anew.

The blood pressure cuff was connected above the ankle at the calf. The level of pressure at which the cuff was inflated varied slightly from *S* to *S*, but a sufficient amount was used to secure a reasonably good wave pattern and still keep the pressure below systolic cutoff point. A 105-mm. level was usually sufficient.

The *S* was requested to assume a relaxed position in an effort to become adjusted to the electrodes and equipment. He was also requested to refrain from extraneous movement of the body, hands, and fingers during each period.

An over-all rest period of 5 min. was observed during which time the skin temperature and skin resistance equipment was rebalanced for optimal recording and the cuff was inflated to obtain a comfortable recording level for *S*.

Design.—The initial period was a 60-sec. recorded rest period for an indication of *S*'s basic level and to note the direction and extent of changes during beginning stages of the experiment. At the end of the 60-sec. period, as at the end of each separate recording period, the pressure of the cuff was released and instruments were turned to zero sensitivity. A 2-min. rest period then occurred.

There were two experimental periods, as two types of stimuli were presented; each followed the same plan in which there was recording of a 30-sec. rest period followed by recording during 23-sec. stimulation. The two types of stress stimuli were: (*a*) assumed "emotion-provoking" —jumping electrical spark via the electric coil; and (*b*) induced muscular tension via hand dynamometer at *S*'s optimal level. The experimental periods were randomized in ABBA order, therefore the complete design was: rest-spark, rest-tension; rest-tension, rest-spark. The *N* was 52, with 26 *S*s for each suborder. The *S*s were psychology students, male and female, from Purdue University.

Recordings.—Our recordings allowed for quantification of skin temperature changes, skin resistance changes, and pulse changes. Pulse

rate was chosen over blood pressure changes because it is believed that greater accuracy can be achieved in recording and counting pulse changes than is possible for blood pressure changes. By tapping arteries accurate blood pressure can be recorded, but this is such a disturbing experience for most individuals that it may mask the influence of the experimental variable to be employed.

Skin temperature and skin resistance measures were in terms of the largest ohm change for the specific period. Base levels were obtained and recorded from rebalancing of the equipment for each subperiod of the record, i.e., rest and stimulus. Centimeter deflections of the stylus movement from the base line, at indicated sensitivity levels, were converted into ohm change values from tables established for each index. The skin temperature quantification differed from skin resistance in that the former was in terms of ohms change, while the latter was in terms of 1000-ohms change.

Changes in skin resistance and skin temperature for the initial rest period were counted from the last 30 sec. for that recording period, as were changes for all other rest periods. Changes during the stimulus periods were on a 23-sec. basis, which was determined by a task of that duration.

Pulse rate changes were always determined by taking a 10-sec. count, and converting the obtained frequency to beats per minute. For the initial rest, the period was the last 10 sec. of the record, as it was for the rest periods. The pulse count was taken from the first 10 sec. of the stimulus periods as no significant differences were found between changes during the first and last 10 sec. of these periods. The pulse change, the final quantification of the pulse index, is the amount of change in pulse rate from the previous subperiod to the period being compared.

Results

The problem of analysis of data was restricted by the type of recordings. The basic interest concerned the relative changes which occurred, and, for the most part, the general question of whether or not the changes were of a similar direction (i.e., increase, decrease, or no activity) within the particular meaning of each index.

Changes in physiological indices.—Inspection of Table 1 reveals that skin temperature changes increased significantly ($p < .01$) over resting states when both spark (assumed emotion-provoking) and tension stimulation were involved. For skin resistance, there were significant decreases during the spark stimulus ($p < .01$) and under the tension stimulation ($p < .01$) as contrasted with changes under resting states. Pulse change also increased significantly ($p < .01$) for both spark and tension stimulus conditions, when compared with resting states.

A comparison of the strength of changes evoked by two types of stimulation is also given in Table 1. Skin temperature increase and skin resistance decrease are shown to have been significantly greater under spark stimulation ($p < .01$) than under

TABLE 1

Differences between Indices under Rest and Stimulation
($N = 52$)

Condition	Skin Temperature†		Skin Resistance‡		Pulse	
	Mean	t	Mean	t	Mean	t
Rest Spark	+ 3.46 −39.63	10.91*	+15.00 −48.83	12.70*	3.23	3.83*
Rest Tension	+ 1.94 −15.86	3.82*	+13.77 −15.48	6.60*	4.73	6.33*
Spark vs. Tension		6.13*		7.53*		1.33

* Significant at the 1% level (2.678).
† Skin temperature changes in ohms change, where minus sign indicates temperature increase.
‡ Skin resistance change in 1000 ohms change, minus sign indicates decrease in resistance.

TABLE 2

CORRELATION RATIOS FOR SKIN TEMPERATURE, SKIN RESISTANCE, AND PULSE DURING REST AND STRESS STIMULATION ($N = 52$)

Condition	Indices Compared: Temperature Versus	Corrected Eta
Initial rest	Resistance	−.219
Rest	Resistance	.487**
Spark stimulation	Resistance	.212
Spark stimulation	Pulse	.286*
Rest	Resistance	.281*
Tension stimulation	Resistance	.458**
Tension stimulation	Pulse	.348*

* Significant at the 5% level (.273).
** Significant at the 1% level (.354).

tension; pulse change increase was slightly larger, but not significantly so, during tension as compared with that for the spark stimulus. These comparisons were made because they might suggest the relative intensity of the stress conditions and because they support a common-sense evaluation.

Relationships between physiological indices.—As scattergrams indicated a tendency away from linearity, correlation ratios were used to determine the degree of relationship existing between the three indices.[1] Table 2 presents the corrected correlation ratios (8). Initial rest periods showed a negative relationship (−.219) which was somewhat short of significance at the 5% level.[2] The rest period just preceding spark stimulation gave a relationship of .478 ($p < .01$), whereas the rest period prior to tension stimulus produced an eta of .281, significant at the 5% level. During spark stimulation, skin temperature and skin resistance showed an insignificant correlation of .212, but temperature and pulse revealed a correlation ratio of .286, significant at the 5% level. Under tension stimulation the relationship between skin temperature and skin resistance was .458 ($p < .01$), while skin temperature and pulse indicated a promising relationship of .381 ($.01 < p < .05$).

DISCUSSION

It seems apparent that under all stress conditions significant skin temperature increases, skin resistance decreases, and pulse increases occurred when compared with reactions recorded during the resting period. It can also be concluded that significant, though not high, relationships were in evidence between changes in skin temperature, skin resistance, and pulse. Skin temperature and skin resistance showed promise of interrelationship. Skin temperature and pulse gave good evidence of some sort of communality. While the neural and circulatory functions believed to produce skin temperature changes seem related to those which gave rise to skin resistance and pulse changes, more research is needed to clarify their interrelationship. It is not clear at present whether skin

[1] Limitations of recording equipment may have cut off complete reactions, and there may have been some buffer action of the body within these limitations. Eta is a more general coefficient than r, in that its application is possible in either linear or nonlinear regressions. Therefore, with at least suspected nonlinearity, eta was used for all comparisons. In all comparisons, the regression for eta was Y on X where Y was either skin resistance or pulse and X was always skin temperature. Thus we find revealed the relation of the change in skin resistance or pulse that resulted with changes in skin temperature. This procedure resulted from the desire to investigate relations between the older, somewhat established indices of skin resistance and pulse and the new index, skin temperature.

[2] Correlation ratios do not have negative signs. However, it is possible to inspect scattergrams in order to substantiate an inverse relation of measures. When such an inspection was made of the scattergrams, only one was found which could possibly be of this direction; that was during initial rest. It was believed that drifting reactions present at this stage accounted for the greater part of this type of relation.

resistance change is influenced by skin temperature change.

The exact relations existing between skin temperature changes and pulse changes remain to be determined. Many investigators discuss the relationship between vascular reactions and temperature changes. The vasoconstriction and vasodilatation of arterioles and their resulting effect on the cutaneous vessels as the basis of decreases and increases in surface temperature have not been satisfactorily demonstrated, although there has been considerable generalization to this effect. Hardy (10) recently discussed vascular responses and indicated that with proper corrections for direct environmental heating and cooling effects, skin temperature could be used as a dependable index of blood flow. Exceptions were apparent, however, since relationships were variable between skin temperature and peripheral blood under changeable conditions. These explanations give some support to the vasoconstriction rationale for skin temperature and blood flow relations, but also support the exceptions cited from Best and Taylor (3) wherein temperature increase may be present without dilatation.

It is believed that neural and general metabolic activities of the organism may have to be considered in relation to the direction and amount of skin temperature change. It appears that many of the statements about the relationship between skin temperature changes and vasomotor reactions have been arrived at by deduction rather than by experimentation.

Our data reveal significant increases in skin temperature under stress stimulation. This, of course, is contradictory to the expected changes when the vasomotor explanation is used. Thus we believe that our experimental evidence, if supported by further investigation, will force a consideration (*a*) of additional factors that influence temperature change, or (*b*) that constriction of the arterioles is not regularly produced through stress.

Summary

Polygraph recordings of skin temperature, skin resistance, and pulse were made of 52 *Ss* exposed to a stress condition assumed to be emotion-provoking and to induced muscular tension.

Under all experimental conditions, significant skin temperature increases, skin resistance decreases, and pulse increases occurred when compared with reactions recorded during the resting period. Several significant, though not high, relationships were found between changes in skin temperature, skin resistance, and pulse.

The findings concerning heat change are in contradiction to the usual statements that stress will produce a constriction of the arterioles and thus bring about a reduction in skin temperature. Skin temperature change is a promising index of emotional reactivity, and may find uses similar to those for GSR and blood pressure.

References

1. Baker, L. M., & Taylor, W. M. An apparatus for recording changes in skin-temperature. *Amer. J. Psychol.*, 1953, **66**, 124–125.
2. Barron, D. H. Vasomotor regulation. In J. F. Fulton (Ed.), *Textbook of physiology*. (16th Ed.) Philadelphia: Saunders, 1949.
3. Best, C. H., & Taylor, N. B. *Physiological basis of medical practice*. (5th Ed.) Baltimore: Williams & Wilkins, 1950.
4. Brobeck, J. Regulation of energy exchange. In J. F. Fulton (Ed.), *Textbook of physiology*. (16th Ed.) Philadelphia: Saunders, 1949.
5. Brown, T. E. The relationship between surface temperature and social traits in young children. *J. genet. Psychol.*, 1939, **55**, 401–413.
6. Courts, F. A. Relationship between experimentally induced muscular tension and memorization. *J. exp. Psychol.*, 1939, **25**, 235–256.
7. Frank, F. E., & Hertzman, A. B. Effects of cigarette smoking on the skin circulation. *Amer. J. Physiol.*, 1940, **129**, 375.
8. Garrett, H. E. *Statistics in psychology and education*. New York: Longmans, Green, 1947.
9. Hardy, J. D. The radiation of heat from the human body. I. An instrument for measuring the radiation and surface temperature of the skin. *J. clin. Invest.*, 1934, **13**, 593–604.

10. Hardy, J. D. Physiological responses to heat and cold. *Annu. Rev. Physiol.*, 1950, **12**, 119–144.
11. Helson, H., & Quantius, L. Changes in skin temperature following intense stimulation. *J. exp. Psychol.*, 1934, **17**, 20–35.
12. Lipkin, M., McDevitt, E., Schwartz, M. S., & Duryee, A. W. On the effects of suggestion in the treatment of vasospastic disorders of the extremities. *Psychosom. Med.*, 1945, **7**, 152–157.
13. Maddock, W. G., & Coller, F. A. Peripheral vasoconstriction by tobacco demonstrated by skin temperature changes. *Proc. Soc. exp. Biol., N. Y.*, 1932, **29**, 487–488.
14. Menzies, R. Conditioned vasomotor responses in human subjects. *J. Psychol.*, 1937, **4**, 75–120.
15. Mittelmann, B., & Wolff, H. G. Affective states and skin temperature: experimental study of subjects with "cold hands" and Raynaud's syndrome. *Psychosom. Med.*, 1939, **1**, 271–292.
16. Mittelmann, B., & Wolff, H. G. Emotions and skin temperature: observations on patients during psychotherapeutic (psychoanalytic) interviews. *Psychosom. Med.*, 1943, **5**, 211–231.
17. Rothe, G. M., McDonald, J. B., & Sheard, C. The effect of smoking cigarettes and the intravenous administration of nicotine on the electrocardiogram, basal metabolic rate, cutaneous temperature, blood pressure, and pulse rate of normal persons. *J. Amer. Med. Ass.*, 1944, **125**, 761–767.
18. Weathery, J. H. Skin temperature changes caused by smoking and other sympathomimetic stimuli. *Amer. Heart J.*, 1942, **24**, 17–30.
19. Ziegler, L. H., & Cash, P. T. A study of the influence of emotions and affects on the surface temperature of the human body. *Amer. J. Psychiat.*, 1938, **95**, 677–696.

(Received March 12, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

THE PERCEPTION OF THE EGOCENTRIC ORIENTATION OF A LINE[1]

IRVIN ROCK

The Graduate Faculty of Political and Social Science, New School for Social Research

In recent years investigation of the problem of visual direction has been more or less restricted to the problem of how we perceive the upright and horizontal directions of space (**1, 2, 3, 4**). The work on this problem has led to the conclusion that there are two factors which account for our accurate perception of the upright direction under varying conditions. On the one hand it has been shown that the main lines of the scene influence the perception of the upright and horizontal directions of space since tilting the axis of the scene produces marked changes in the direction of the phenomenal upright. The phenomenal upright is displaced in the direction of the upright of the tilted scene. On the other hand it has been shown that the direction of the upright is still perceived fairly accurately even where no visual scene is present at all, as for example in a completely darkened room (**5**). It is true that if *O* is tilted in this situation, judgments of the upright are no longer quite so accurate. Nevertheless, within a certain range of error, it is still correct to say that we can perceive the general direction of the upright even under these circumstances. From this it has been concluded that the second factor is the direction of the force of gravity on the body. Ordinarily it is presumed that these two factors cooperate with each other since the direction of gravity is aligned with the vertical lines of the visual scene.

It is doubtful, however, whether these two factors tell the whole story about perception of the upright. In the situation where the visual scene is eliminated, the orientation of the image of the line whose orientation is being judged cannot in any way be determined in relation to the image of a frame of reference as is possible when a visual scene is present. Nevertheless it must be kept in mind that *O* is performing a visual task in that the judgment is based upon a visual stimulus. The problem then is, how, with the postural cues provided by a particular tilt of the body, a particular retinal orientation becomes the proximal stimulus for perception of the upright. Regardless of how this problem is eventually solved, the fact remains that the orientation of the retinal image must be one component of the stimulus situation underlying perception of the upright.

However, in addition to speaking of the orientation of a line with respect to gravity—i.e., upright, tilted, or horizontal in space—introspective examination of daily experience suggests that we may also speak of a line's phenomenal orientation with respect to the self. In this sense one might refer to the line as appearing to be parallel to the sagittal axis of the head, tilted with respect to it, etc. We have no experimental knowledge of the

[1] This paper is based on a portion of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Division of Philosophy and Psychology of the Graduate Faculty of Political and Social Science of the New School for Social Research. The writer wishes to express his appreciation to Professors Hans Wallach and Mary Henle for their many helpful suggestions throughout the course of this investigation.

stimulus conditions which might underlie this type of experience. It is clear, however, that, as in the case of perception of the upright, such an experience must be based upon a visual stimulus, namely, the orientation of the line.

The question then arises whether—as in the case of perception of the upright—the experience of a line's orientation to the self must also depend in some way upon postural cues of the direction of the force of gravity. Since, if *O* is tilted from the upright, the direction which is upright in space (i.e., the environmental upright direction) is no longer the same as the one which is parallel to the sagittal axis of the head (i.e., the sagittal anatomical direction), it is clear that the retinal orientation which defines the former is not always the same as that which defines the latter. Nevertheless, one might suppose that perceived orientation with respect to the self also depends upon postural cues of the direction of gravity. Hence one might ask the question: Is perception of orientation with respect to the self possible without reference to postural cues as to the direction of gravity?

It is entirely possible that in the case of phenomenal orientation with respect to the self, the orientation of the retinal image alone, without any additional cues provided by the direction of gravity, is a sufficient proximal stimulus. This means that particular retinal orientations determine particular phenomenal orientations of lines with respect to the self regardless of *O*'s orientation with respect to gravity. Thus a line stimulating a sagittal retinal orientation, for example, may always be experienced as parallel to the sagittal axis of the head regardless of the position of *O* in space. For this to be established, however, it would have to be shown that the orientation of the image can mediate such perceived orientation even when the possible role of gravity has been eliminated.

In other words, although for direction with respect to the upright of space (perception of the environmental upright) both orientation of retinal image and posturally given direction of gravity are important, it may be hypothesized that for orientation of a line with respect to the self (perceived egocentric orientation), only the former is necessary. Since, in daily life, the head changes its orientation while the upright remains fixed, there is no one retinal orientation which is the constant proximal stimulus for the perception of the upright. We continue to perceive the upright accurately when we ourselves are tilted although environmentally upright lines now fall upon tilted retinal orientations (different retinal orientations for different degrees of tilt). This is why perception of the environmental upright must be mediated by the combined action of the orientation of the image and the posturally given direction of gravity as was pointed out above (leaving aside for the moment the important role of the main lines of the visual scene in situations where one is present). On the other hand, regardless of the orientation of *O* with respect to gravity or in any other respect, a line which remains parallel to, say, the sagittal axis of the head will continue to stimulate a sagittal retinal orientation. The eyes do not change their orientation to the head to any appreciable extent.[2] Hence

[2] Some changes of eye orientation with respect to the head occur because of cyclotorsion of the eyes. Also some changes occur in such unusual cases as the counterrolling of the eyes when *O* is inclined laterally. It remains to be investigated whether in such cases the egocentric orientations of lines are actually displaced.

there very well can be a retinal orientation which is the proximal stimulus for a particular experienced egocentric orientation.

The reason we speak of orientation with respect to the head, rather than with respect to the entire body, is that it is really only appropriate to speak of a possible correspondence of retinal orientation with perceived orientation of a line with respect to the *head*. A given retinal orientation can maintain its egocentric directional significance only with respect to the head simply because the eyes do not change their orientation to the head to any appreciable extent, as mentioned above. Since the head may very considerably change its orientation to the trunk, obviously no fixed relation between retinal orientation and trunk is to be expected.

It is the purpose of the experiment to be described to demonstrate that the egocentric orientation of a line can be perceived under conditions where postural cues of the direction of gravity cannot possibly play a role. Furthermore, by eliminating eye movements and other factors as possible determinants, the experiment is intended to demonstrate the dependence of egocentric orientation on retinal orientation and to make a preliminary exploration of the exact nature of this relationship. It, perhaps, should be added that the use of the term "retinal orientation" in this paper is not meant to imply anything about the anatomical location of the neural correlate of perceived egocentric orientation. "Retinal-cortical orientation" would probably be more accurate in this respect, but we are only concerned here in referring to the proximal stimulus conditions in a purely descriptive sense.

Method

The basic plan of the experiment was to require *O* to judge the orientation of a luminous line with respect to the sagittal and left-right axes of his head while lying on his back on a table (supine position) in a totally dark room. The line rotated about a vertical axis in a horizontal plane 11 in. above the level of *O*'s eyes (see Fig. 1).

Under these circumstances it may be concluded that gravity does not play any role in the judgments. The line remains in the same relation to the force of gravity at all times, namely, in a horizontal plane; hence, no one position of the line would be in any way favored with respect to it. The possibility of touch or kinesthesis influencing the judgments in any other way was eliminated by preventing *O* from lifting his arms from the table. Furthermore, since the room was completely dark, it may be concluded that the judgments were not influenced by any visual frame of reference. As a further precaution, *O* was blindfolded before entering the experimental room, disoriented after entering it, and thus brought into the supine position on the table without any knowledge whatsoever as to where he was lying with respect to the walls of the room. Hence, it cannot be argued that anything like a frame of reference represented in memory can have contributed to the outcome.

It was also decided to eliminate the possibility that the judgments were based upon or improved upon by cues from eye movements. If *O* were allowed unrestricted eye movements during the time of observation of the line, he might sweep his eyes along its length thereby providing himself with a possible cue as to the orientation of the line with respect to his head. It was not necessary to eliminate eye movements in order to prove that the egocentric orientation of a line can be perceived independently of any role of gravity since, of course, the possible kinesthetic cues based on eye movements should not be confused with kinesthetic cues having to do with gravity. As noted above, however, another purpose of this experiment was to show that perceived egocentric orientation depends upon the orientation of the retinal image and not upon any other stimulus conditions.

The following procedure was, therefore, adopted. The line could be made to flash on for a brief fraction of a second (approximately .2 sec. in duration). Prior to each exposure of the line, *O* was given a signal to fixate a small luminously painted spot ($\frac{1}{4}$ in. diameter) at the center of the line. Shortly after *O* fixated, the line was exposed. Thus, although eye movements of appreciable extent perhaps *can* occur

in less than the time of exposure of the line, it is felt that the additional feature of fixation makes such a possibility unlikely. It may also be pointed out that even if occasionally *O*'s eye did move during exposure of the line it is still unlikely that this could account for his over-all average score of many settings.

Finally, *O* was allowed to use only one eye, the other eye remaining blindfolded throughout the experiment. It is possible that binocular vision might provide cues to the orientation of the line: a double image of the line produced by far or near fixation might be a cue because the width between the images would be maximum for the egocentric sagittal and zero for the egocentric left-right direction of the line.

A word should be said in advance about how the results are to be interpreted. The hypothesis that the egocentric orientation of a line can be perceived under the conditions of this experiment does not set any exact criterion to be met for its acceptance or rejection. What the hypothesis really amounts to is the contention that *O*s will be able to accomplish this to the extent that their judgments of the direction of the line will not simply be random in the sense that any one direction would be as apt to be chosen as any other. In other words it is by no means necessary and in fact it is impossible to predict the degree of accuracy with which this task can be accomplished.

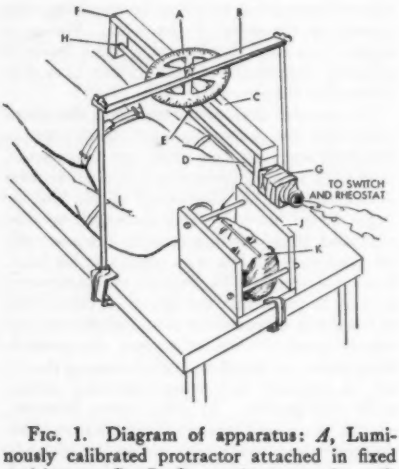

FIG. 1. Diagram of apparatus: *A*, Luminously calibrated protractor attached in fixed position to *B*; *B*, Supporting cross bar; *C*, Luminous marking for reading angular position of *F*; *D*, Lucite rod, ¼ in. in diameter and 14 in. long; *E*, Luminously painted fixation spot on underside of *D*; *F*, Lucite rod support; *G*, Wooden block containing small incandescent bulb covered with black friction tape; *H*, Silver-paper capping on end of lucite rod (hidden from view in figure); *J*, Headrest lined with foam rubber; *K*, Dowel taped along sagittal dividing line of *O*'s face serving as "zero" position.

Apparatus.—A drawing of the apparatus is shown in Fig. 1. A lucite rod is illuminated by a small incandescent bulb placed inside the wooden block so that the light can only escape through the opening in which one end of the rod is snugly fit. The other end of the rod is capped by silver paper so as to reflect back a maximum of light which illuminates the rod fairly homogeneously along its length. The bulb is connected with a telegraph key. A tap of the key turns on the light for about .2 sec. The lucite rod is directly underneath and parallel to the wooden supporting rod so that the position of the latter as indicated by the protractor reading also gives the position of the former. The intensity of the light is adjustable by a rheostat.

Procedure.—The procedure was explained to *O* outside the experimental room before he was blindfolded. He was told that he would be required to judge when a luminous line is parallel to either the long (sagittal) or the short (left-right) axis of his head; this was made fully clear by demonstrating with a pencil exactly what was meant by these two directions. Each eye was then separately blindfolded so that one eye could be uncovered when *O* was in position and the other could remain covered throughout. The *O* was led into the room, disoriented, and brought into the supine position on the table. His head was fixed firmly in the headrest, and a small dowel was taped along the sagittal dividing line of *O*'s face (see *K* in Fig. 1). This remained throughout the experiment and served as the guiding line for *E* in determining the zero or correct direction of the luminous rod for the sagittal settings (and therefore indirectly for the left-right settings). The lucite rod was brought to the zero protractor reading and the position of *O*'s head was then adjusted until the lucite rod was exactly parallel to the dowel. A small plumb line was used to bring the eye to be used directly underneath the luminous fixation spot. The headrest was then clamped to the table.

The room lights were then turned out and the blindfold was removed from one eye. The *O* was told to close the eye between each trial. At the signal "ready," he was to open his eye and look off to the side; at the signal "fix" he was to fixate the luminous spot and observe the direction of the line when it appeared.

The luminous rod was exposed at successive 5° intervals until *O* indicated its position to be correct. He was allowed to request *E* to show

TABLE 1

MEAN SETTINGS AND *SD*'S OF THE SAGITTAL AND LEFT-RIGHT EGOCENTRIC ORIENTATIONS

Observer		Settings to Equal Sagittal			Settings to Equal Left-Right		
		Mean	*SD*	σ_M	Mean	*SD*	σ_M
Left Eye							
Males	1	1.5° CW*	5.0°	1.7°	7.0° CCW	12.2°	4.1°
	2	2.0 CW	2.4	.8	1.5 CCW	2.3	.8
	3	4.0 CW	5.8	1.9	4.0 CW	4.4	1.5
Females	4	4.0 CW	2.0	.7	8.5 CW	3.8	1.3
	5	1.5 CCW	3.2	1.1	6.0 CCW	6.6	2.2
	6	4.5 CCW	6.5	2.2	10.0 CCW	7.4	2.5
Right Eye							
Males	7	9.0 CW	3.7	1.2	9.0 CW	3.0	1.0
	8	3.0 CW	3.3	1.1	1.5 CW	2.3	.8
	9	5.0 CW	5.5	1.8	8.5 CCW	7.8	2.6
Females	10	5.0 CW	2.2	.7	2.0 CW	2.4	.8
	11	8.5 CW	3.2	1.1	4.0 CCW	5.7	1.9
	12	3.5 CCW	5.5	1.8	14.0 CCW	8.0	2.7
Mean		2.7° CW	4.0°		2.2° CCW	5.5°	
SD		4.0	1.5		7.0	2.9	

* CW = Clockwise; CCW = Counterclockwise.

the rod again in a certain position or to reverse the direction in which the rod was being rotated. Trials for the left-right and sagittal directions were alternated; the rod was moved alternately in clockwise and counter-clockwise directions on trials for each of these directions. Altogether ten sagittal and ten left-right settings were obtained for each *O*. In order to prevent *O* from judging according to the constant number of intervals from the starting position the latter was varied from trial to trial anywhere from 85° to 0° from the correct position with the bulk of starting positions falling somewhere between 40° to 60° from it.

Subjects.—There were 12 *O*s varying in age from 16 to 34 yr. Six were males and six were females. Six *O*s used their left eye and six their right eye. There was no need for *O*s to be naive as to the purpose or procedure of the experiment. None, however, knew the position of the table in the room. The majority had never been in the room before.

RESULTS

The results are given in Table 1. The mean setting of each *O* was computed as an algebraic mean of the clockwise or counter-clockwise displacements of the line when it was judged to be "sagittal" ("left-right"), and is a measure equivalent to a constant error. The *SD*'s were computed from the obtained mean settings.

All *O*s were able to judge the egocentric orientation of the rod with considerable but by no means perfect accuracy and consistency. The mean settings range from 1.5° to 9.0° (clockwise or counter-clockwise) for the sagittal and from 1.5° to 14.0° (clockwise or counter-clockwise) for the left-right orientation. The *SD*'s range from 2.0° to 6.5° for the sagittal and from 2.3° to 12.2° for the left-right settings with mean *SD*'s of 4.0° and 5.5°, respectively. This means that retinal orientation does—within the limits of this accuracy and consistency—determine the egocentric orientation of a visual line. It may be mentioned in passing that in a preliminary experiment with seven *O*s, each given four trials for the sagittal

orientation only and allowed unrestricted eye movements, the results were quite comparable. The mean settings ranged from .25° to 7.75°. These similar results substantiate the assumption that eye movements do not play an important role in this task.

It should be pointed out that there is a certain limitation to the accuracy of measurements obtained in this experiment. The head is not a perfectly symmetrical geometric shape so that there are no precisely defined sagittal and left-right lines which may be taken as the zero positions for the settings. In the opinion of *E* these zero directions can only be given within a tolerance of one or two degrees.[3] Hence the settings obtained must be regarded as accurate measures of deviation from the sagittal and left-right axes of the head only within this degree of tolerance; settings deviating by this amount or less can then be taken to be nondeviating settings, whereas those deviating by more than this amount can be regarded as indications of a real difference between phenomenal and objective orientation with respect to the head.

The mean of the 12 *O* means for the sagittal settings is 2.7° clockwise (CW) and for the left-right settings is 2.2° counter-clockwise (CCW). These values are small and are not significantly greater than zero. Hence there is no evidence that there is a direction preference or constant error in the population as a whole.

On the other hand we see in Table 1 that, in several individual cases, there was a tendency to judge the rod as parallel when it was actually displaced in one particular direction. This is evidenced by the fact that for several *O*s the mean setting for the sagittal orientation or left-right orientation or sometimes both is significantly greater than zero in either the CW or CCW direction. Also in most of these cases the mean setting is sufficiently greater than 2° so that it must be regarded as really displaced from the sagittal or left-right axis of the head. This result would be understandable if we assume that the retinal orientation serving as the proximal stimulus for the phenomenal sagittal (left-right) orientation is, for many *O*s, either momentarily or permanently displaced from the objective sagittal (left-right) retinal direction in one particular direction. This interpretation makes particular sense in cases like *O* No. 7, where the mean setting for the left-right orientation is displaced in the same direction as the mean setting for the sagittal orientation and by the same amount, because one would logically expect that if one of the phenomenal directions is displaced the other would also be displaced so as to preserve the right angle relationship between them.

As noted above, half the *O*s were males and half females and half used the left eye and half the right eye. The purpose of this was simply to provide a sample which would not reflect any special conditions or type of *O*s. Nevertheless these factors might be responsible for differences in results although the number of cases in these groups is too small to warrant statistical comparisons. It might be supposed that the use of the left or right eye could influence the direction of settings. The results here were as follows: For the sagittal settings, the mean of the six means was .9° CW for the left eye and 4.5° CW for the right eye. For the left-right settings the mean of the six means was 2.0° CCW for the left eye and 2.3° CCW for the right eye. Thus the direction of error is the same for the two eyes in both cases.

In the case of sex, it might be supposed that there could be some difference in ability to perform this task. The most suitable measure of ability in this case is consistency (*SD*) since the mean settings may reflect individual constant errors and not inaccuracy as such. The mean of the *SD*'s for the six males was 4.3° and 3.8° for the six females for the sagittal settings. For the left-right settings these values were 5.3° and 5.7°, respectively. Hence the consistency of performance was, within the limits of these

[3] It might be argued that a line parallel to the direction which passes through the center of the pupils of the two eyes would serve to define precisely the *left-right* orientation. This however need not always be perpendicular to that direction which appears to bisect the face in the sagittal orientation.

small samples, about the same for maes land females.

Finally, the question arises as to whether or not there was any difference in difficulty in making the sagittal or the left-right settings. In this case there was evidence that *O*s found it more difficult to perform the latter task. This can be seen by the fact that, as determined by a *t* test, the mean of the 12 *SD*'s for the left-right trials is significantly higher than the mean of the 12 *SD*'s for the sagittal trials at the .05 level of confidence. Furthermore, whereas of the total of 120 settings made by the 12 *O*s there were only three of the sagittal orientation which deviated by as much as 15° from the zero position, there were 12 settings of the left-right orientation with this degree of error. Four of these 12 deviated by 20° and one by 25°, whereas none of the three "sagittal" cases exceeded 15°. There were also qualitative signs of the greater difficulty of the left-right trials.

Discussion

The results confirm the assumption of the dependency of perceived egocentric orientation on retinal orientation but also show that this holds only within certain limits of accuracy. The mean settings for many *O*s do not precisely coincide with the sagittal and left-right axes of the head. Furthermore, the size of the *SD* points up the fact that for a given *O* there is a range of retinal orientations within which any one may be the proximal stimulus for a particular phenomenal orientation of a line at a given moment (sometimes there is a considerable range of uncertainty for a given *O* even during a given trial). Some *O*s showed a range of 25° or more from their extreme CW to their extreme CCW settings. Nevertheless, since as noted above, the hypothesis only demanded settings which were better than random, it can certainly be regarded as having been substantiated.

To understand these results it is important to be clear about the following: The fact that the relationship of points or lines to one another in the outside scene is preserved in the retinal image (and in the pattern of excitation in the visual cortex) only means that we should expect these relationships to be preserved in visual experience. But this does not demand any correspondence between orientation of retinal stimulation and perceived orientation of a line with respect to the self. We should expect to experience an angle of 90° as a right angle because the angular relationship of lines to each other is preserved in the retinal image. But this does not logically require that this angle be experienced as in any particular orientation with respect to the self. That one of the sides of the angle happens to be parallel to the head and therefore stimulates a sagittal retinal orientation does not require us to expect, on logical grounds, that it will be experienced as parallel to the sagittal axis of the head and not, say, tilted with respect to it. The objective correspondence of a line's orientation to the head with its image's orientation to the head does not logically require a correspondence between the phenomenal orientation of the line to the head with its image's orientation to the head.

Our hypothesis that such a correspondence would nevertheless be found was therefore not based on any of the facts now known about retinal projection or the cortical representation of the retinal image. Rather, it was based on introspective observation and, as noted earlier, the fact that at least there could be a constant proximal stimulus direction for perceived egocentric orientation because of the more or less unchanging relationship of eye to head orientation. Our results show that while a fairly close correspondence between retinal orientation and phenomenal orientation to the head obtains it is not a precise correspondence. This bears out the logical consideration discussed above because by contrast the correspondence between the objective and the phenomenal orientation of two lines to each other is precise. Small differences in the orientation of a line relative to another line are readily perceived. Apparently orientation with respect to the self is not as finely discriminable.

The fact that in our experiment the orientation of the line is characterized with respect to the egocentric axes of *O*'s head should not be confused with the

question of the location of objects in space with respect to *O*. The location of objects from *O* is also given in egocentric terms, namely, left, right, above, below, but here these terms refer to the actual location, radially, from *O* (or from his mid-eye position as origin) of the object. The egocentric orientation of a line as studied in this experiment, however, refers to its orientation within a frontal-parallel plane with respect to the egocentric axes of the head and not, of course, with where the line, as a whole, is located with respect to *O*. This distinction does not mean that these two problems may not turn out to be related in certain respects.

Finally, it should be pointed out that although this experiment has demonstrated that egocentric orientation is perceived independently of postural cues of the direction of gravity, this does not mean that under other circumstances such cues may not play some role. It is possible that perception of egocentric orientation would be more accurate or at least would be changed in some way in a situation where the line rotates about a horizontal axis so as to change its orientation with respect to the direction of gravity. Hence this problem requires further exploration.

Summary

The purpose of this experiment was to test the possibility that the egocentric orientation of a line (its apparent orientation with respect to the head) can be perceived under conditions where postural cues of the direction of gravity cannot possibly play a role. This was accomplished by requiring *O*s to judge the orientation of a luminous line with respect to the sagittal and left-right axes of the head while lying in a supine position on a table in a totally dark room.

By eliminating other cues such as eye movements it was possible to show that the perception of egocentric orientation depends upon the orientation of the retinal image of the line. It was found that *O*s could accomplish this task with fairly good but by no means perfect accuracy or consistency. The study thus demonstrates the correctness of the hypothesis and presents factual data concerning how well such a task is actually performed.

References

1. Asch, S. E., & Witkin, H. A. Studies in space orientation: I. Perception of the upright with displaced visual fields. *J. exp. Psychol.*, 1948, **38**, 325–337.
2. Asch, S. E., & Witkin, H. A. Studies in space orientation: II. Perception of the upright with displaced visual fields and with body tilted. *J. exp. Psychol.*, 1948, **38**, 455–477.
3. Gibson, J. J., & Mowrer, O. H. Determinants of the perceived vertical and horizontal. *Psychol. Rev.*, 1938, **45**, 300–323.
4. Wertheimer, M. Experimentelle Studien über das Sehen von Bewegung. *Z. Psychol.*, 1912, **61**, 161–265.
5. Witkin, H. A., & Asch, S. E. Studies in space orientation: III. Perception of the upright in the absence of a visual field. *J. exp. Psychol.*, 1948, **38**, 603–614.

(Received February 15, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

THE EFFECT OF COMPONENT PRACTICE ON PERFORMANCE OF A LEVER-POSITIONING SKILL[1]

GEORGE E. BRIGGS[2] AND W. J. BROGDEN

University of Wisconsin

The problem of whole-part learning has been treated almost entirely in the context of acquisition of relatively large amounts of verbal materials. If part learning is defined in terms of the behavior components involved in a motor task, a problem similar to that of whole-part learning of verbal materials occurs. What is the relative efficiency in acquisition of a skill when practice is given on the several components separately and on the whole task? There are special problems of part learning of skill. The fractionation of a motor task will probably not result in homogeneous parts. Differences in difficulty of acquisition for the several components should be expected as the rule. Components will differ in terms of movements required and may differ also in terms of muscle groups and body members involved. The movement components of many tasks are spatially and temporally symmetrical. Others are asymmetrical. In some tasks, a particular component may be dependent upon the prior occurrence of another component. These factors suggest that whole-part learning of motor skill is quite different and perhaps considerably more complex than whole-part learning of verbal materials. In any case, a relatively simple experimental approach to whole-part learning of skill appears desirable. The present study used a lever-positioning task for which the right-left and the forward-backward dimensions of movement may be taken as separate components of the whole task. Four practice schedules were designed to provide an evaluation of whole-part practice and of the relative difficulty of the two components of the task.

[1] This research was supported in part by the United States Air Force under Contract Number AF 33(038)-23294 monitored by the Armament Systems Personnel Research Laboratory, AF Personnel and Training Research Center, Lowry Air Force Base, Denver, Colo. Reproduction in whole or in part permitted for any purpose of the United States Government.

[2] Now postdoctoral fellow of the National Science Foundation at Northwestern University.

METHOD

Apparatus.—A modified U. S. Air Force Complex Coordination Test, Model E, (4) was employed in the present study. The modification consists of immobilization of the rudder foot controls and screening of the lower horizontal row of lights related to the rudder. The display of the modified apparatus thus consists of two double rows of lights, one in the horizontal and one in the vertical plane. Each row is made up of 13 pairs of red (stimulus) and green (response) lights. Position of the lever in the right-left dimension of movement controls the green response lights of the horizontal display, and position in the forward-backward dimension of movement controls the green response lights of the vertical display. Thirteen random combinations of the vertical and horizontal red stimulus lights are presented serially by means of a stepping switch. The *S* must position the lever so that a match is held for .5 sec. in order to change the display. Daily calibrations of the timing circuit were necessary in order to hold constant the .5 sec. duration throughout the experiment.

Subjects.—Thirty-two male students at the University of Wisconsin were employed as *S*s in this study. None had had previous experience with the apparatus. Each *S* served for ten 30-min. sessions and was paid $8.50 for participation in the experiment.

Procedure.—The *S* was seated in the apparatus and given instructions for the task (4).

On the first experimental session, he was given ten 2-min. trials of whole practice during which both the vertical and horizontal stimulus lights had to be matched before the display pattern was changed. These initial standard trials were given to all *S*s, and performance at this session was used to equate the four groups. Following the initial session there were seven practice periods for which the nature of practice varied by group. For Group 1 the eight *S*s were maintained on the standard whole practice with ten trials per session. For Group 2 the eight *S*s were divided equally into two subgroups, Groups 2a and 2b. The four *S*s in Group 2a first received one 2-min. standard trial followed by four trials during which only the backward-forward component was practiced. The next four trials involved practice on the right-left component only. The final trial was one of standard practice. The four *S*s of Group 2b received the same treatment as did the *S*s of Group 2a except that practice on the horizontal component preceded practice on the vertical component. The eight *S*s of Group 3 received one standard trial followed by eight trials on the backward-forward component. Each session ended with one standard trial. For Group 4 the eight *S*s had one initial standard trial of practice, eight trials of practice on the right-left component, and one final standard trial.

At the conclusion of this treatment, the *S*s in all four groups received two daily sessions of ten 2-min. trials of standard practice on the whole task. During every practice session a 1-min. rest interval occurred between Trials 1 and 2, a 2-min. rest interval occurred between Trials 5 and 6, and a 1-min. rest interval was given between Trials 9 and 10. A 15-sec. interval occurred between all other trials.

The *E* recorded the number of matches made by each *S* at the end of each 2-min. trial. These measures represent the basic data of the experiment. The score at the end of each trial was given to *S* in order that he might have immediate knowledge of results.

Results

The primary results of the experiment are presented in Fig. 1. The mean performance of the whole task on Trials 1 and 10 of each practice session is plotted for each of the four groups. Examination of these curves shows that although substantial learning occurs for all groups, whole practice provides for acquisition superior to that of any of the part methods. The part methods do not appear to differ among themselves except for the possibility of superior performance on the part of Group 2

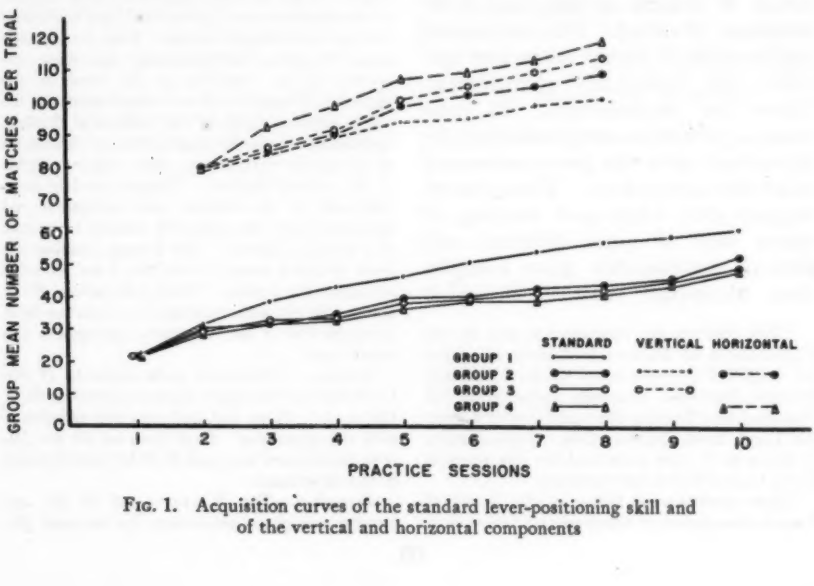

Fig. 1. Acquisition curves of the standard lever-positioning skill and of the vertical and horizontal components

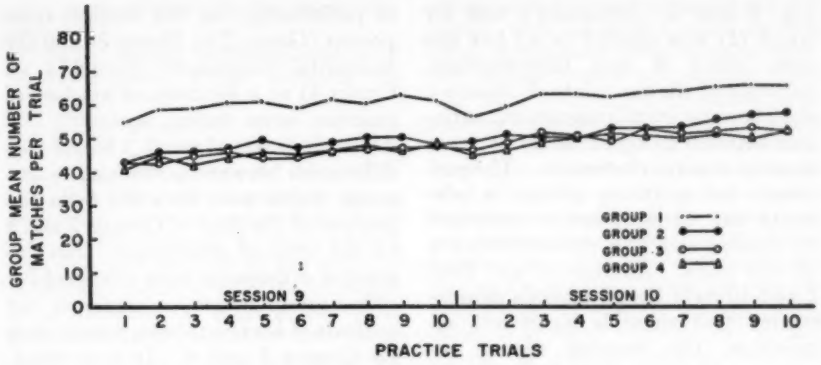

FIG. 2. Terminal performance on the standard lever-positioning skill

on the final practice session. Because of this, and because of the possibility that Groups 2, 3, and 4 may have increased their rate of acquisition on the standard practice provided by Sessions 9 and 10, mean group performance on each one of the ten practice trials of these two sessions is plotted in Fig. 2. Rate of acquisition on the part of Groups 2, 3, and 4, appears to be very similar and closely parallel to that for Group 1. The exception to this generalization previously noted for Group 2 makes its appearance on Trials 8, 9, and 10 of Session 10. The experiment did not provide sufficient terminal practice on the whole task to determine if and when the performance of Groups 2, 3, or 4 would equal the performance of Group 1.

The four upper curves of Fig. 1 are plots of the mean performance on component practice by Groups 2, 3, and 4 for Sessions 2–7. Group 2 received four trials of the vertical component and four trials on the horizontal component in each session. Groups 3 and 4 received eight trials on the vertical component and horizontal component, respectively. Rate of acquisition appears to be greater for components than for the whole task. Performance in terms of number of matches per trial of component practice appears to be somewhat greater than twice that for whole practice. This ratio varies, however, since performance on the vertical component is consistently less than that on the horizontal component, and mean performance based on eight trials of practice per session is greater than that based on four trials per session. Thus, there appears to be a difference in difficulty of the vertical and horizontal components of the lever-positioning task. The difference between Group 2 and Group 3 on the vertical component, and between Group 2 and Group 4 on the horizontal component, however, may be a function of interference between alternation of component practice on the part of Group 2 as well as a function of the difference in absolute number of trials that exist between Group 2 and Groups 3 and 4.

A series of statistical tests of the data was completed in order to test the reliability of the observations noted above under the discussion of

Fig. 1 and 2. Alexander's test for trend (1) was applied in all but one case, when it was inappropriate. Analysis of the ten trials of Session 1 shows no statistically significant variation between groups in terms of either slope or mean performance. The procedure for matching groups is adequate and performance as measured by Trials 1 and 10 is representative of all the trials. Analysis of the Trial 1 and 10 data for all sessions demonstrates that there is significant acquisition, the learning curves are nonlinear, and the nonlinearity of the learning curves is different for the different groups. The latter result represents the difference between the curve for Group 1 on the one hand and the curves for Groups 2, 3, and 4 on the other and supports the superiority of the whole method of practice over the part methods in the acquisition of the skill. Analysis of the data for Sessions 9 and 10, the curves of Fig. 2, shows that there is significant acquisition for each group, the curves are nonlinear, and the curve for Group 1 is different from the curves for Groups 2, 3, and 4 taken together. The latter conclusion is supported by application of the Duncan range test (2) to the group means.

Analysis of performance on component practice, the upper curves of Fig. 1, relate to differences between performance on the vertical and horizontal components of the task at two levels of practice. An initial analysis of the Group 2 data was performed by extending the analysis of variance scheme presented by Edwards (3, pp. 288–296). There is a significant trend in the data over the seven practice sessions, indicating learning. Neither the comparison between Subgroups 2a and 2b nor that between the horizontal and vertical components was significant. The differences in level of performance on the vertical component (Group 2 vs. Group 3) and the horizontal component (Group 2 vs. Group 4) as a function of amount of practice were tested separately by Alexander's test of trend. Significant differences between group means and group slopes were obtained from the analysis of the data of Groups 2 and 3 on the vertical component. No significant differences were obtained for either group means or slopes in the analysis of horizontal component data for Groups 2 and 4. It was noted, however, that individual variability on the horizontal component was about twice that of the individual variability on the vertical component. This may account for the unreliability of the group differences in performance of the horizontal component.

Discussion

The results of the present experiment clearly indicate the superiority of whole practice over the part methods in the acquisition of a lever-positioning skill. There are no reliable differences in acquisition between practice on components with the part methods for either measures of the whole performance on Trials 1 and 10 of all practice sessions or on all trials of the final two practice sessions. The complexity of the whole-part problem in the acquisition of motor skill seriously limits the generality of the results of any single experiment. One might, however, expect results similar to those of the present study for whole-part practice of other two-dimensional tasks. There is the possibility that the difference between the whole and part methods may decrease or disappear with more extensive practice than that involved in the present experimental design. The fact that performance on the whole task was still improving at the termination of practice and that rate of improvement on component performance appears to be greater than that on the whole perform-

ance gives support to the above suggestion.

A more striking theoretical possibility occurs, however, if one considers the relation between the effect of whole-part practice on acquisition to level of task complexity. If complexity is defined as being positively related to difficulty of acquisition by the whole method and to number of fractional components of the task, the advantage of the whole method over the part methods at low levels of complexity might be expected to change to no difference at some greater complexity to a difference in favor of one or more of the part methods at a still greater complexity. A serious problem at the higher complexity levels may occur because of heterogeneity of the components in terms of relative difficulty. With many components of a motor task, relative difficulty might vary considerably. Part practice involving only the most difficult component or components may provide for more efficient acquisition of the skill than will whole practice or part practice on the relatively less difficult components. The task studied in the present experiment is one of relatively low complexity and the two components appear to be homogeneous in difficulty. There is, however, an unreliable numerical superiority of the horizontal component over the vertical component. Although part practice on these two components separately produced no significant difference in performance of the whole task, an increase in the amount of practice may produce a differential effect. Amount of part practice produced a significant difference in level of performance in the components, but no test of this effect on the whole performance was included in the experiment. Both the total amount of practice and the amount of component practice may be critical parameters of the effect of whole-part practice on acquisition of motor skills and consideration of them must be given in experimental studies.

Measurement of the effect of part or component practice on the whole performance may be made either by means of test trials on the whole performance introduced into each component practice session with terminal practice on the whole performance or by terminal practice on the whole performance only. The latter method will provide a purer measure of transfer, and hence a purer measure of the effect of part practice. In the former method, the measurement of whole practice on each practice session provides for practice on the whole task, and the terminal measures on the whole task represent the effect of both part and whole practice. Since the two are confounded, no separate measure of the effect of the part methods can be obtained. It may be best to treat this procedure as a method of part practice and to measure the effect of all part methods by means of transfer to performance on the whole task. This would require the elimination of test trials of the whole method in individual practice sessions from most of the part method procedures. If test trials of whole practice are used, another base for measuring the effect of component practice would be provided by a group that receives only the two trials of whole practice with a rest in between. Comparison of the results for such a group with those of the groups given component practice would show whether component practice produces a positive or negative transfer effect on performance of the whole task. Even though this refinement of experimental design would provide for more precise measurement of the effect of component practice, there would still be the confounding of part and whole practice for the experimental groups. Consideration of these problems also raises a question about the use of initial practice on the whole method. Although this provides a means of matching groups in initial ability, it is possible that this practice has an effect on subsequent part practice and on terminal whole practice. An experimental solution of this problem also appears to be desirable.

Viewed within this theoretical framework, solution of the whole-part problem in the acquisition of motor skill will require the development of a scale of

complexity based upon difficulty of acquisition by the whole method of practice and upon the number and nature of fractional components of the task. The development of such a scale will require the use of tasks that are relatively extensive in time, since multiple component fractionation of a task will depend upon its temporal as well as upon its spatial characteristics. Consideration must also be given to further refinement of the methods of part practice and of measuring the effect of component practice on the whole task. The total amount of practice and the relative amount of practice by the part methods must be treated as critical parameters in the relation between whole-part practice and acquisition of the whole skill.

Summary

Four groups of eight *Ss* each practiced a lever-positioning skill on ten practice sessions. Group 1 received ten 2-min. trials of the complete task on each of the ten sessions. The *Ss* of the other three groups practiced the whole performance on all trials of Sessions 1, 9, and 10 and on Trials 1 and 10 of all other practice sessions. Group 2 received four trials of practice on the horizontal component and four trials on the vertical component for Sessions 2 through 8. Group 3 received eight trials of practice on the vertical component only and Group 4 eight trials on the horizontal component only for Sessions 2 through 8.

There was a clear superiority of whole practice over part practice. No reliable differences in performance were obtained between the three groups that received part practice. For all groups, however, there was significant acquisition over the practice sessions. All the learning curves showed significant departure from linearity. Performance on the components was slightly more than twice that on the whole task. The numerical superiority of Groups 3 and 4 on component practice over that for the subgroups of Group 2 is significant only for the vertical component. In the case of performance on the horizontal component, however, individual variability was about twice that of individual variability on the vertical component. In discussing the results, it is suggested that the superiority of whole practice may disappear at high levels of task complexity.

References

1. Alexander, H. A. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.
2. Duncan, R. B. A significance test for differences between ranked treatments in an analysis of variance. *Virginia J. Sci.*, 1951, **2** (new series), 171–189.
3. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
4. Melton, A. W. *Apparatus tests.* Washington, D. C.: U. S. Gov't. Print. Off., 1947. (*AAF Aviat. Psychol. Program Res. Rep.* No. 4.)

(Received for early publication March 29, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

MULTIPLE RESPONSE TRANSFER AS A FUNCTION OF SUPPLEMENTARY TRAINING WITH VERBAL SCHEMATIC AIDS [1]

FREDERICK H. KRESSE, ROBERT M. PETERSON, AND DAVID A. GRANT

University of Wisconsin

Most evidence on transfer training indicates that the maximum transfer is obtained when the training task is identical with the transfer task. In many industrial and military situations, however, operators are trained to use one set of equipment and subsequently are shifted to equipment of somewhat different design. Aside from the practical problem, it is of theoretical interest to consider what kinds of training procedures would maximize transfer to situations where important stimulus or response components have been changed. It was conjectured that transfer from one type of equipment to another would be facilitated by verbal conceptualization of the task if the basic function of the original and transferred equipment was similar. This conjecture was based on Judd's observation that superior transfer of training was in some cases obtained when the principle involved in the task had been taught verbally to *S* (**4**). The present study was aimed at testing this conjecture in the context of a complex psychomotor task.

The task in the present study involved operating the hand control of a simulated radar antenna in such a way as to aim the simulated antenna at a hypothetical target aircraft. Two radar displays were used, a B display with range on the vertical axis and azimuth on the horizontal axis, and a C display with elevation on the vertical axis and azimuth on the horizontal axis. The *S*s were trained on the B or C display and transferred to the C or B display, respectively, with and without supporting verbal descriptions and conceptualizations of the functions of the simulated radar. It was expected that the verbalized conceptualizations of the radar task would assist in transferring from B to C display and vice versa.

Prior to the present study an experiment was carried out on a smaller scale (40 *S*s) using a similar trifactorial experimental design but with some differences in instructions and procedure.[2] The conceptually instructed groups proved generally inferior to those instructed to respond mechanically, and this inferiority was statistically significant on transfer tests but not during initial training. The hypothesis of the present study appeared to have been disproved. Close examination of the data, however, disclosed that the conceptual groups had higher error variance than the mechanical groups particularly on the initial transfer trial, and the conceptual groups showed markedly inferior performance on the initial training trial followed by extraordinary improvement on the second and succeeding training trials. This suggested strongly that after the initial training trial many of the conceptually instructed *S*s had departed from their instructions with resulting "improvement" in training performance but serious confusion in

[1] This research was supported in part by the United States Air Force under Contract AF 33(038)-23294, monitored by the Armament Systems Personnel Research Laboratory, Air Force Personnel and Training Research Center, Lowry Air Force Base, Denver, Colorado. Permission is granted for reproduction, translation, publication, use, and disposition, in whole and in part, by or for the United States Government.

[2] A report of the earlier research, emphasizing display comparisons, was submitted as a working memorandum to the Aero-Medical Laboratory of the Wright Air Development Center. It was entitled "A Preliminary Comparison of the B-Display and C-Display with respect to Target Acquisition Time."

the transfer tests. This is a common problem when *E* attempts to manipulate one of *S*'s responses (such as conceptualizing targets in space) as the independent variable of an experiment. It was felt that the conceptualization hypothesis deserved a more adequate test with safeguards to insure that *S* had adequate practice in conceptualizing the target in space. To accomplish this the instructions were improved and an additional paper and pencil procedure was introduced. A detailed analysis of the data of the present study indicated that these procedures eliminated the idiosyncrasies noted in the data of the earlier study.

Method

Apparatus.—The apparatus was designed to provide an approximate simulation of radar procedures, but no particular radar was simulated. There were four independent units in the apparatus: the display unit, the control stick assembly, the programming unit, and the error counters.

The display unit was an RCA oscilloscope placed about 3 ft. in front of *S* at eye level. The scope face was divided into four quadrants by a plastic overlay. Three overlays were used with the different experimental procedures: one was of green plastic with unlabeled axes; another was of clear plastic divided into four quadrants with the word "range" on the left-hand vertical scale and "azimuth" on the bottom horizontal scale; the third was of clear plastic divided into four quadrants with the word "elevation" on the vertical scale and "azimuth" on the horizontal scale. Each target pattern appeared in the center of one of the four quadrants. Responses made by *S* changed the display in characteristic ways so that *S* was presented with a continuous feedback of response information.

The control-stick assembly was mounted on a pedestal placed between *S*'s legs and could be reached and operated comfortably from a sitting position. The stick itself was about 9 in. high and was fitted with a plastic hand grip. A movement of 1.5° activated switches to alter the display. On top of the grip was a small thumb control which could be moved either forward or backward.

The programming unit consisted of a Western Union Tape Transmitter which was used to determine the patterns that would appear on the oscilloscope display unit. There were also associated relays, and signal generators, to produce the patterns on the oscilloscope, and response detection circuits which balanced all output voltages when the correct response was made and which activated the solenoid that advanced

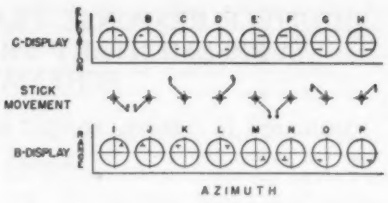

Fig. 1. Diagrams of display patterns and corresponding hand-control responses for the C display and B display

the transmitter to the next pattern. The programming unit operated on a self-paced basis so that it advanced as soon as a correct response was made.

An error occurred whenever the control stick or thumb control was moved 1.5° in the wrong direction. The error counters were connected to the programming unit and indicated separately the number of errors in azimuth, elevation, and range responses. A buzzer sounded each time *S* made an error. In Fig. 1 are shown the eight possible display patterns used with the C display and the eight possible patterns for the B display. When a pattern appeared on the oscilloscope, *S* was to make the appropriate control-stick and thumb movements which are indicated in the diagrams between the C display and B display patterns. The displays are as seen in the vertical plane, but the stick movements are as seen from above the horizontal plane. For example, C display Pattern F consists of a long horizontal line in the upper left-hand quadrant. The correct response involves moving the control stick backward to the left, and moving the thumb control back, as indicated by the arrow, so that the pattern is centered in the scope and reduced to medium length; the next pattern is then presented. The corresponding B display pattern would be N which consists of a horizontal line in the lower left-hand quadrant with a vertical pip above the line. The pip represents the elevation of the target. The control-stick and thumb-control movements to correct this pattern are the same as for Pattern F. In the C display, the width of the horizontal line corresponds to target distance. A long line represents short-range targets and a short line represents long-range targets. Incorrect motions of the control stick on Target F would move the target further to the left, or raise it higher on the scope, or make it appear longer. Incorrect motions on Target N would lengthen the vertical pip, lower the target on the scope face, or move it further to the left.

Experimental design.—The experiment used a trifactorial design. The three factors were the display (B or C), the instructions (mechanical or conceptual), and the transfer condition (same or different displays during training and transfer test). The differences between the displays have been explained. In the transfer tests half of the *S*s were tested on the same display on which they were trained, and the other half of the *S*s switched displays for the test series. The instructions differed in that half of the *S*s were instructed as to how the movements of the control stick and thumb control altered the target display and were instructed to make the correct responses as quickly as possible so as to center each target display and obtain the next pattern. The other half of the *S*s were instructed to conceptualize the situation. They were told that each pattern represented a target that was either far or near in *range*, to the right or left in *azimuth*, and either above or below in *elevation*. Whereas the mechanically instructed *S*s used the unlabeled, green overlay, the conceptually instructed *S*s used the appropriate clear, plastic overlays.

Subjects.—Eighty male undergraduates were used as *S*s. They were randomly assigned to one of the eight groups with the limitation that all groups were filled concurrently. All *S*s were paid for their services.

Procedure.—When each *S* appeared for the experiment, he was first read a set of instructions appropriate to the particular procedure to which he had been assigned. Mechanical instructions were straightforward. The conceptual instructions emphasized that the apparatus simulated a radar display, and that the pattern on the oscilloscope represented a target aircraft. Their task was to move the hand control so as to point the radar antenna directly at the target and to "range in" or "range out" depending on the indicated range of the simulated target. This would result in centering the target on the scope face. The *S*s were instructed always to think in terms of the location of the target aircraft in the three-dimensional space as they centered their antenna on the target. Each *S* was then given two sample patterns and made the two correct responses in order to familiarize himself with the operation of the equipment.

On the first day, each *S* was given five trials consisting of 48 patterns each. Between each practice trial, *S* was given a series of 24 paper and pencil problems. On the second day, the same procedure was followed, and after the last training trial there was a 5-min. rest period during which *S*s were given additional instructions if they were to transfer to a different display. There followed five successive test trials of 48 patterns each.

The paper and pencil task consisted of a series of diagrams on which *S* had to indicate the correct stick responses. In the mechanically instructed groups, *S* observed the pattern on the oscilloscope, marked the correct stick responses on the paper and pencil test, and then pressed a button to change to the next stimulus pattern. In the conceptually instructed groups, *S* looked at the stimulus display, marked the spatial position of the target on a diagram, and then marked another diagram indicating the correct stick movements before pressing the button to change to the next pattern. The conceptual groups were thus forced to locate the targets in space before marking the stick movements. The display was the same as in the regular training trials, but the response was by means of paper and pencil. Thus, there was no feedback from the display regarding the adequacy of response as was the case in the regular training trials. As soon as the twenty-fourth diagram was marked, *S* pushed aside his writing board, and *E* started the next performance series on the apparatus.

Results

Response time.—Performance on this psychomotor task was scored in terms of time per stimulus pattern, whether it was on the apparatus or on the paper and pencil task. Errors were counted separately for range, elevation, and azimuth.

All of the performance data using the complete apparatus are presented graphically in Fig. 2 where time per stimulus pattern is plotted against successive trials during practice and during the transfer test. In Fig. 2, it is apparent that during the practice trials performance on the C display is consistently superior to performance on the B display. The relationship between the conceptual and mechanical groups is not consistent for the B and C displays throughout the practice trials.

The statistical analysis of the trends in the practice data of Fig. 2 is given in Table 1. Here it is seen that there are significant differences between the displays with respect to the linear, quadratic, and higher-order

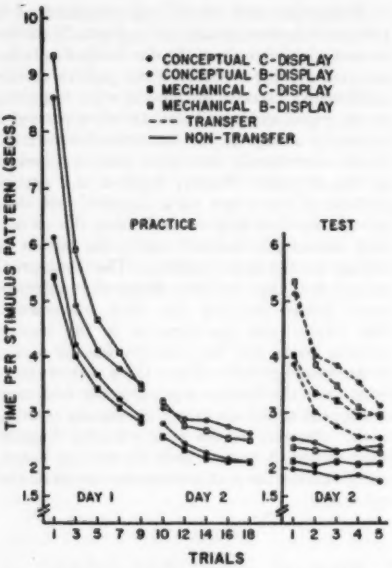

FIG. 2. Learning curves and transfer test data. Mean response time per pattern is plotted as a function of trials during practice and transfer test. Data for odd-number trials only are given during training. The paper and pencil task occupied even-numbered trials.

components of the trends. The significance of these components indicates that the learning curves for the B display show greater slope and greater curvature than the learning curves for the C display. There is also a significant linear component in the interaction between instructions and displays. This interaction indicates that the mechanical B display curve and the conceptual C display curve show greater slopes, respectively, than the conceptual B display and mechanical C display curve.

Turning to the transfer test data of Fig. 2, it is apparent that all transfer groups show a high initial performance loss, but the loss is not great enough to reduce their performance to the initial level of Day 1. It is noteworthy that the transfer groups never attain the control group levels in performance during the transfer tests. Transferring from the C display to the B display results in the greatest loss and the greatest recovery over the five test trials. The data of the transfer tests are analyzed in Table 2. In Table 2, it is apparent that there are significant differences due to the displays, both on the initial transfer test trials, on the final two transfer trials, and on recovery from Trial 1 to Trial 5 in the transfer test series. The transferred groups are different from the nontransferred groups on all of these measures. The only other significant components of the analysis are the Display × Transfer interaction for the first test trial and the recovery from the first to the fifth test trial. Here there is some indication that the difference between the B and C displays in the nontransferred group is not as great as the difference between the B and C displays in the transferred groups. In other words, transferring to the B display enhances its relative difficulty. When the data of the transferred groups, only, were investigated, it was discovered that the only significant difference among them was due to the displays. Performance on the B display was significantly inferior to performance on the C display for the transferred groups.

The evaluation of all transfer results must take into account the pretransfer levels of performance. Since the B and C displays were of unequal difficulty it was possible experimentally to equalize amount of practice or level of performance but not both. Theoretical and practical grounds favored equalizing practice trials so that at the end of pretransfer training there were apparent differences in performance levels (Fig. 2 and 3) usually favoring the C display. These differences are

TABLE 1

SUMMARIES OF EXTENDED† ALEXANDER TREND ANALYSES OF TIME SCORES DURING TRAINING TRIALS

Source	Error Term Row No.	*df*	*F*'s		
			Instructions	Displays	Instructions × Displays
1. Between Group Slopes	8	1	1.13	32.88***	6.76*
2. Between Group Quadratics	9	1	0.01	31.11***	1.26
3. Between Group Deviations from Linear and Quadratic Estimation	10	7	1.17	10.81***	0.33
4. Between Group Means	11	1	0.00	25.20***	1.18

Source	*df*	Mean Square	*F* (Error term in Row 10)
5. Over-all Slope	1	1273.83	4226.26***
6. Over-all Quadratic	1	330.16	1095.39***
7. Over-all Deviations from Linear and Quadratic Estimation	7	18.55	61.56***
8. Between Individual Slopes	76	1.98	6.56***
9. Between Individual Quadratics	76	1.27	4.22***
10. Between Individual Deviations from Linear and Quadratic Estimation (*Error Term*)	532	*0.301*	—
11. Between Individual Means	76	6.52	21.62***

† The Alexander Trend Test (1) was extended to deal with quadratic components of trend and the orthogonal factors, *Instructions* and *Displays*.
* Significant at 5% confidence level.
*** Significant at .1% confidence level.

probably not to be attributed to practice per se, however, because they existed on the initial training trials. If the pretransfer differences were taken into account, by analysis of covariance or use of difference scores, the differences favoring transfer from B to C displays would, except on azimuth errors, be increased artifactually so that it was decided to

TABLE 2

SUMMARIES OF ANALYSES OF TRANSFER TEST TIME SCORES

Source	*df*	A. Score on First Test Trial	B. Average Score on Transfer Trials 4 and 5	C. Difference in Scores between Trials 1 and 5
1. Displays	1	18.62***	9.24**	8.47**
2. Instructions	1	.20	.26	.00
3. Transfer	1	142.90***	34.92***	87.48***
4. Display × Instructions	1	.10	3.58	1.33
5. Display × Transfer	1	5.21*	.01	8.76**
6. Instructions × Transfer	1	.22	.08	.42
7. Display × Instructions × Transfer	1	.21	.57	.70
8. *Error Variance Estimates*	72	*.7668*	*.3703*	*.5876*

* Significant at 5% confidence level.
** Significant at 1% confidence level.
*** Significant at .1% confidence level.

restrict statistical analysis to the unadjusted transfer test scores.

Errors.—The error data for the experiment are shown in Fig. 3. Here the average number of errors per stimulus pattern is plotted separately for range, elevation, and azimuth against successive trials during practice and transfer tests. With respect to range error, it is apparent that all groups performed approximately equally well during the training trials. There were no statistically significant differences on the last training trial. On the transfer tests there was a tremendous performance loss, especially in transferring from the C to the B display. Transfer from the B display to the C display procedure was also poor for the mechanically instructed groups. At the same time, transfer from the B display to the C display was excellent in the conceptual group. The poor performance of the conceptual B display group that was not transferred on the test series was due to one *S*. The statistical analysis of the range errors on the transfer test series is given in Section A of Table 3. Here it is seen that the differences due to transferring are significant throughout the transfer test series, and that the display difference on the first transfer trial is significant as well as

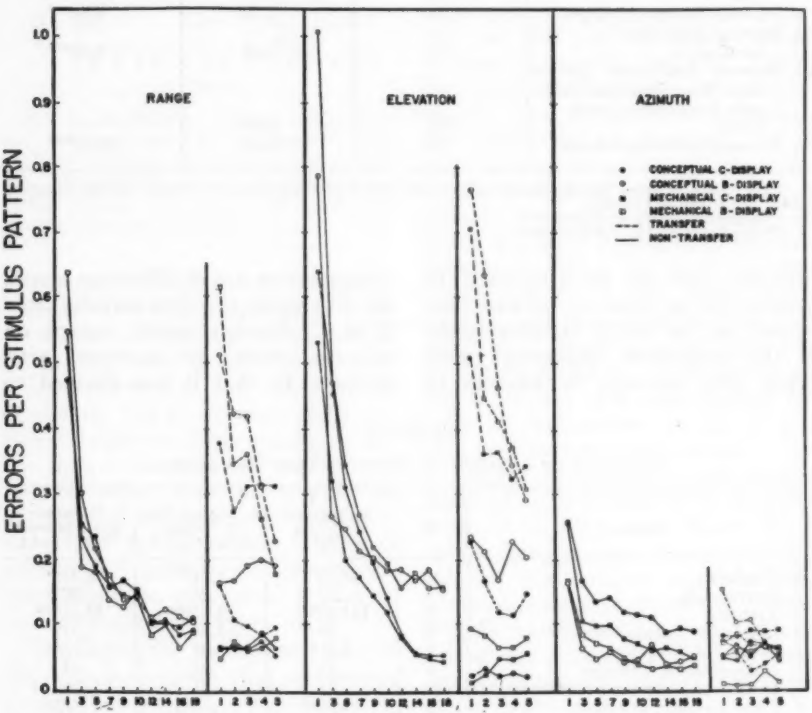

FIG. 3. Learning and transfer test curves for error data

TABLE 3

SUMMARIES OF ANALYSES OF ERROR SCORES ON THE TRANSFER TEST

Source	*df*	A. Range Errors			B. Elevation Errors			C. Azimuth Errors		
		First Transfer Trial	Average Trials 4 and 5	Difference Trials 1–5	First Transfer Trial	Average Trials 4 and 5	Difference Trials 1–5	First Transfer Trial	Average Trials 4 and 5	Difference Trials 1–5
1. Displays	1	12.89***	2.19	12.69***	19.81***	5.69*	23.27***	1.26	1.01	6.43*
2. Instructions	1	1.13	.17	.85	.11	.40	.10	3.28	.01	3.68
3. Transfer	1	50.29***	11.38***	38.12***	64.01***	19.75***	57.03***	6.88*	.75	2.70
4. D × I	1	1.49	7.92**	1.19	4.22*	4.28*	.79	7.21**	3.04	1.20
5. D × T	1	7.19**	.01	19.99***	3.65	0.02	13.47***	2.64	.23	3.68
6. I × T	1	5.42*	3.92	1.32	2.57	2.82	.24	.04	1.57	.57
7. D × I × T	1	.00	.09	2.66	.65	1.02	1.71	.66	.00	1.36
8. *Error Variance Estimate*	72	*1.7672*	*1.0531*	*1.0772*	*2.6731*	*1.4641*	*1.0973*	*.2226*	*.1746*	*.1849*

* Significant at 5% confidence level.
** Significant at 1% confidence level.
*** Significant at .1% confidence level.

the degree of recovery of the transferred groups. Instructions introduced no differences in performance save that they played a role in interaction with the transfer variable on the first transfer test trial, where conceptual instructions resulted in superior transfer as compared with the mechanical instructions. This F of 5.42 was significant at only the 5% level and was the only direct evidence favoring the hypothesis that conceptual instructions generally facilitate transfer.

On elevation errors, the C-display performance was superior to the B-display performance throughout the practice trials, and this difference was significant at the 1% level ($F = 9.63$) on the final practice trial. There was a great loss in transferring from the C display to the B display, and somewhat less loss in transferring from the B display to the C display. Transfer from the B display to the C display was especially good for the conceptually instructed groups. In general, the picture on elevation errors was the same as that on range errors. The statistical analysis of the elevation errors is given in Section B of Table 3 where the display and transfer effects are found to be statistically significant in all transfer test variables that were analyzed. There was a Display × Instruction interaction on the initial test trial and on the last two test trials. This interaction was caused primarily by the relatively high degree of transfer from the B to the C display in the conceptual group. This is an added indication that conceptual instructions facilitate transfer from the B to C display. The poor transfer in range and elevation errors was due to the fact that the vertical axis on the B display was range and the vertical axis on the C display was elevation, so that switching from one display to the other interchanged the meaning of the vertical dimension of the displays. The result was virtually no positive transfer relative to initial training scores, in range performance and very little positive transfer in elevation performance.

The azimuth errors showed a different picture from the range and elevation errors. During training trials, the performance on the B display was superior to performance on the C display, although the difference was not statistically significant on the last training trial. There was a significant

Display × Instruction interaction ($F = 4.07$), as the C display differences between conceptual and mechanica instructions were greater than the B display differences. There was clear-cut positive transfer in spite of the consistent loss in performance in all the transferred groups. Transferring from the B to the C display resulted in slightly better performance in azimuth errors. In the statistical analysis of the azimuth errors in Section C of Table 3, it is seen that the only significant display difference was the recovery from Test Trial 1 to Test Trial 5. Switching displays (*Transfer*) made a significant difference on Trial 1, only. There was a significant Displays × Instructions interaction on the first test trial. The superior performance throughout training and transfer on the azimuth component of the displays was probably due to good stimulus response compatibility (2). The relatively high degree of positive transfer was due to the fact that on both the B and C displays azimuth occupied the horizontal axis. (It should be pointed out that the stick movements were identical relative to the *target position* in space on both the B and C displays for range and elevation, as well as azimuth.) The lack of 100% transfer was due to the change between the display and stick relationship in range and elevation. This change momentarily affected the azimuth performance where all S-R relationships were constant.

Paper and pencil task.—Performance on the paper and pencil task showed very nice learning curves through the eight trials on Days 1 and 2. On the paper and pencil task, performance on the C display was consistently superior in time per diagram to performance on the B display. The relative superiority of the C display was greater in the mechanically instructed groups than in the conceptually instructed groups. These relations held throughout the training period.

Discussion

It was observed that the conceptual groups showed better response time performance on the transfer tests than did the mechanically instructed groups. In terms of the Gagné-Foster index (3), there was a 55% to 57% transfer in the groups with conceptual instructions as opposed to 43% transfer in the mechanically instructed groups that transferred to the C display, and 57% transfer in the mechanically instructed groups that transferred to the B display. These differences, however, were not statistically significant. The only statistically significant differences due to instructions were in the range and elevation error scores for the conceptually instructed groups that transferred from the B to C displays. Conceptual instructions did not apparently facilitate transfer from the C to B display. On the whole, it may be said that the basic hypothesis of the experiment, viz., that conceptualizing the spatial relationships simulated in the apparatus would facilitate transfer from one display to another, was not *generally* verified, although the evidence indicated that the conceptual instructions facilitated transfer from the B to the C display.

The conceptual instructions given in the present experiment were perhaps less central to the performance than was the case in the historical experiment of Judd (4). In a sense, the conceptually trained *S*s of the present study had to learn something over and above the actual performance. There was some indication on the preliminary experiment mentioned in the introduction that *S*s were confused or hindered by the conceptual instructions at the outset of practice. Those *S*s apparently gave up the attempt at conceptualization and devised systems of their own, so that on the test series, when the conceptual groups were told that they should use

their conceptual system to interpret the new task, it presented almost a new problem to them. The improved instructions and the paper and pencil task of the present experiment greatly improved the transfer performance of the conceptual group, and it is noteworthy that the additional task of localizing the simulated target in space did not actually decrease the over-all speed of performance in the conceptual groups as compared with the mechanical groups in the final experiment.

Some special discussion must be directed toward the peculiar nature of the independent variable of the present study. Usually experimenters manipulate stimulus variables, but in this instance an attempt was made to manipulate *S*'s response, spatial conceptualization, in such a way as to affect the dependent variable, transfer. In attempting indirectly to manipulate *S*'s conceptual set the stimulus conditions of instructions, oscilloscope overlay, and the paper and pencil procedures were manipulated together. These three were tied together as one over-all procedural complex, and their individual effects are impossible to separate. There is some indication, comparing the results of the preliminary and present experiments, that all three were needed to induce a stable, effective set. In any case further exploration will be required to isolate the effects of each of the three parts of the procedural complex. An unfortunate minor confounding arose because the conceptual paper and pencil procedure took a little longer than the mechanical paper and pencil procedure. The conceptual groups therefore received somewhat greater distribution of training on the principal task. The times amounted to an average of about 190 sec. as compared with about 170 sec. for each set of 24 diagrams. Neither training nor transfer data show any clear-cut effects of this distribution variable, and no positive effects of distribution of practice on transfer of training have been reported elsewhere, so that the influence of this factor can probably be discounted in the present results.

In addition to its apparent contribution to the differences between the preliminary and final experiments, practice on the paper and pencil task probably transferred to *S*s' performance on the apparatus. When the time per pattern curves of the preliminary experiment and the final experiment were superimposed on a graph which plotted performance data as a function of number of trials on the apparatus exclusively (i.e., ignoring the paper and pencil trials) the performance of *S*s in the final experiment was clearly superior to that of *S*s in the earlier experiment during the later stages of practice. This suggests that the paper and pencil task provided effective practice for transfer to the apparatus task. The amount of transfer from paper and pencil performance to apparatus performance is to be investigated in supplementary studies.

The findings on error performance are easily explained. There were great increases in errors in elevation and range because these display dimensions had been interchanged in the transfer tests when *S*s had to change from B to C displays, and vice versa. There were relatively fewer errors in azimuth and a high degree of positive transfer. Azimuth was coded the same way on both B and C displays.

The difference in performance on the B and C displays is of practical interest. The earlier experiment showed insignificant differences favoring the C display, but the final experiment showed a clear-cut difference favoring the C display, both in training and in transfer tests. Although the differences were not great, they were statistically significant. This is consistent with the widely held impression that the C display is more naturally coded than the B display and that there is greater stimulus response compatibility (2) in the C display than is the case with the B display.

Summary

The purpose of the experiment was to test the hypothesis that verbal conceptualization would facilitate transfer of training from one

method of coding on a display to another when the data coded (range, azimuth, and elevation of a simulated target) are the same.

The task required *S* to make a three-component stick response to a stimulus pattern presented on an oscilloscope. Mechanically instructed groups were told what responses to make to change the pattern's location and character in a prescribed way, whereas the conceptually instructed groups were given paper and pencil training in referring the patterns to hypothetical target aircraft in different spatial positions as shown on a radar scope. The stick movements corresponded to moving the radar antenna left or right, up or down, and ranging in or out, so as to get "on target." Control-stick movements were constant relative to the spatial position of the simulated target, regardless of differences in coding on the two types of displays.

Simulated radar B and C displays were used. The 80 *S*s were run in a trifactorial design where display, instructions, and transfer were the three factors. Transferred *S*s switched displays after ten 48-pattern, self-paced, training trials; nontransferred *S*s continued to work on the same display. The transfer test consisted of five 48-pattern trials.

The principal findings were as follows:

1. During practice trials, the C display was superior on all performance measures except azimuth errors.

2. Response time and error measures revealed some positive transfer from C to B displays, but transfer was highest from B to C displays.

3. Conceptualization did not facilitate transfer from C to B displays, but both range and elevation error scores showed that conceptualization facilitated transfer from B to C displays.

4. On the basis of poor results with conceptual instructions alone in a preliminary experiment using 40 *S*s in the same design, the conceptual instructions were improved, and a paper and pencil task which provided drill in conceptualizing target position was interspersed with the training trials. Comparison of the results of the two experiments suggested that the paper and pencil task effectively supplemented the conceptual instructions and also provided some direct transfer to the apparatus performance.

References

1. Alexander, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.
2. Fitts, P. M., & Seeger, C. M. S-R compatibility: spatial characteristics of stimulus and response codes. *J. exp. Psychol.*, 1953, **46**, 199–210.
3. Gagné, R. M., Foster, H., & Crowley, M. E. The measurement of transfer of training. *Psychol. Bull.*, 1948, **45**, 97–130.
4. Judd, C. H. The relation of special training to general intelligence. *Educ. Rev.*, 1908, **36**, 28–42.

(Received for early publication June 7, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

TWO- AND THREE-DIMENSIONAL AUTOKINETIC MOVEMENT AS A FUNCTION OF SIZE AND BRIGHTNESS OF STIMULI[1]

WARD EDWARDS

The Johns Hopkins University[2]

It is customary to think of autokinetic movement as an illusion in which a small, dim, objectively stationary spot of light is seen as moving. But large, bright stimuli might perhaps work just as well. Indeed, examination of the literature shows that large stimuli have been successfully used. Bourdon (**2**) was able to see a little movement of a stimulus 25 cm. in diameter at 2-m. distance (14° visual angle). Adams (**1**) used a 10° stimulus for incidental observations and also did a systematic experiment on the effects of sizes up to 1° 26′. Recently Karwoski, Redner, and Wood (**7**) did a systematic experiment in which they used various large stimuli, and varied both size and brightness. In their main experiment, they used a structure shaped like a cross with the center omitted. The whole structure subtended 11° (except in a few incidental observations from a shorter distance, where it subtended 26°). Their *Ss* reported movement so long as the stimulus was not too bright. They found, as usual, that the parts of the composite figure usually moved together, as a unit. Similarly, Luchins (**8, 9**) found that a 1° 42′ stimulus was not seen as moving if the brightness was high enough, with different brightnesses necessary to stop movement for different *Ss*. Various recent authors (e.g., **6**) have used line figures of various sizes and shapes as stimuli, and have reported that autokinetic movement was seen.

A persistent technical difficulty in research on autokinetic movement of large, bright stimuli is that as soon as a stimulus is made large and bright, it gives off a great deal of light. This light illuminates other objects in the room, and so they become visible too. They provide a "frame of reference" for the autokinetic stimulus, and so no movement is seen.

It seems quite possible that if the autokinetic stimulus could be prevented from illuminating its surroundings, then quite large, quite bright stimuli might be seen as autokinetically moving. This experiment was designed to find out whether this is so, and to find out whether size and brightness of stimulus affect the latency and amount of movement seen for those *Ss* who do see the large, bright stimuli as moving.

When an autokinetic stimulus is observed, the most striking thing seen is movement in the frontal-parallel plane. This kind of autokinetic movement has been most extensively investigated, and it has become customary to use methods of reporting direction of movement which permit

[1] This experiment was done under Contract N5-ori-166, Task Order 1, between the Office of Naval Research and The Johns Hopkins University. This is report No. 166-I-183, Project Designation No. NR 145-089, under that contract. I am indebted to Miss Frances Brown, who ran the *Ss* and processed most of the data, and whose alertness made possible the discovery of movement in the third dimension in the first experiment. I am also indebted to Mr. A. E. Snider, who designed the apparatus.

[2] Now at the Armament Systems Personnel Research Laboratory, Air Force Personnel and Training Research Center.

only two-dimensional reports. The main experiment reported here was designed in the conventional way, which makes reports of movement in the third dimension very difficult, but spontaneous reports of three-dimensional movement were made by several *S*s, and so the whole question of autokinetic movement in the third dimension was explored with two auxiliary experiments.

Experiment I

Method

Apparatus.—The basic apparatus of this experiment was a room, shown in Fig. 1. Light was produced by seven 60-w. Lumiline tubes behind a ground glass screen in the light box. The voltage applied to the tubes was controlled to vary brightness. Masks slid in front of the ground glass controlled the size and shape of the stimuli. The two partitions divided the room into three subrooms. All surfaces in all three subrooms were painted flat black. In the two partitions were set thin aluminum plates, and the holes through which *S* looked were cut in these plates. This reduced the width of the edge of the partition, and so reduced its visibility. These aluminum plates were painted flat black, and in addition they were kept smoked with candle smoke. These measures reduced the amount of light reflected by the edges of the partitions to below threshold at all brightness levels used in the experiment. When *S* sat with his head on the chin rest, he could see only the stimulus defined by the mask in front of the ground glass screen. The apertures in the partitions were large enough so that he could not see their edges even if he moved a bit, so long as he kept his chin on the chin rest.

All these measures were designed to prevent illumination of other objects by the large, bright stimuli. But there is one set of objects with which these measures do not deal: the nose, cheeks, and eyebrows of *S*. To reduce reflection from these, they were painted with theatrical black face make-up. This reduced the reflection substantially, but not to zero. Any other way of removing these stimuli from the visual field would have introduced other more undesirable stimuli.

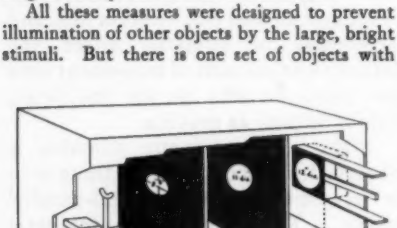

Fig. 1. The apparatus

All presentations of autokinetic stimuli were accompanied by the sound of an electric motor used as a sound effect.

Subjects.—The 51 *S*s were male students. Most of them were Johns Hopkins undergraduates, but a few were students from local high schools. All were naive about autokinetic movement and about experimental psychology in general.

Stimuli.—Ten masks were used, and each mask was exposed at three brightness levels. The brightness levels used, as measured by a Macbeth illuminometer located at the place where *S* observed and looking at the largest stimulus, were .0927, 3.85, and 173.9 ft.-L. The dimmest stimulus was very orange in color; the brightest appeared white. The brightest stimulus appeared about as bright as a page of white paper under an exceptionally good reading light. The masks were five circles and five equilateral triangles. The diameters of the circles and the sides of the equilateral triangles measured .25, 1, 4, 8, and 12 in.

Procedure.—Upon arrival, *S* was placed in his chair and dark adapted for about 10 min. (more adaptation would have been useless because the bright stimuli would have destroyed the adaptation). He was then told that he would be shown a series of moving lights, and that failure to see their amount and direction of movement correctly would be evidence of neuroticism. He was provided with a button, which he was instructed to press when he saw movement start. After the stimulus went out, he reported the direction of movement by stating the number on an imaginary clockface toward which the movement seemed to go. He was then shown a 12-in. luminous line created by putting an appropriate mask in front of the light box. No *S* reported autokinetic movement of this line, perhaps because the motor sound effect was not used with it and no one expected it to move. The *S*s were asked to report the amount of movement they had seen as a percentage of the length of this line. This arrangement prevented changes in how long *S* thinks an inch is. Then the next stimulus was presented.

The five circles, each presented at three brightnesses, make 15 stimuli. Two groups of 15 *S*s each were run on these 15 stimuli in a latin-square order. Thus the design is a replicated latin square. The remaining 21 of

TABLE 1

NUMBER OF FAILURES TO SEE MOVEMENT ON FIRST PRESENTATION OF CIRCLES FOR 21 REJECTED *Ss* IN EXP. I

Brightness	Size in Inches					Total
	.25	1	4	8	12	
Low	3	1	4	9	9	26
Medium	1	4	9	11	15	40
High	4	5	9	13	13	44
Total	8	10	22	33	37	110

the 51 *Ss* were rejected, for reasons explained below. After a rest period, the 15 triangular stimuli were presented to each *S* in the same order in which he had seen the circular stimuli.

When *S* pressed his button, it started a timer which turned off the stimulus 5 sec. later. If he failed to press it 60 sec. after the stimulus was turned on, it was turned off again and the trial was recorded as a failure. The same stimulus was presented again later. If *S* failed to see movement on it three times, he was rejected from the experiment.

Results

Rejected subjects.—Twenty-one of the 51 *Ss* were rejected for failure to see movement in three trials on one or more stimuli. It is important to ask whether these rejections are correlated with the major variables of the experiment, size and brightness. Table 1 shows the number of failures to see movement of a circular stimulus on its first presentation. The number of failures increased monotonically with increasing size and brightness. However, it is not true that the failures occur only on the large, bright stimuli. Instead, the pattern of failures parallels very closely the pattern, which will be reported below, of increases in latency and decreases in amount for the successful *Ss*. If all *Ss* were included in Table 1, instead of only rejected *Ss*, it would not change the picture; the 30 accepted *Ss* failed on the first presentation of a circular stimulus a total of only three times.

Latency and amount.—For each stimulus, a latency before *S* pressed his key to report movement, an amount of movement seen, and a direction of movement were recorded. Table 2 shows the mean amounts and latencies of movement as functions of size and brightness for the 30 successful *Ss*. It is apparent that the directions of the effects are just what was expected and analogous to what was found for the rejected *Ss*: latency increases and amount decreases with increasing size and brightness. The only exception is the mean amount of movement seen on the 4-in. circle, which is less than it should be to preserve a monotonic relationship between size and amount. The functions, though monotonic, are not linear. Analyses of variance showed that all these main effects are highly significant, except for the effect of brightness on the amount of movement of triangles. The effects of brightness are in general somewhat less than the effects of size. The changes in amount of movement seen are substantially smaller and con-

TABLE 2

MEAN LATENCIES AND AMOUNTS OF MOVEMENT SEEN IN EXP. I

Variable	Latencies in Seconds		Amounts in Inches	
	Circles	Triangles	Circles	Triangles
Size in inches				
12	15.2	15.8	3.46	2.61
8	13.1	14.6	3.75	2.87
4	10.1	9.8	3.04	3.19
1	7.5	8.1	3.94	3.81
¼	7.1	7.2	4.13	3.96
Brightness				
High	12.4	12.7	3.27	3.13
Medium	11.5	11.6	3.45	3.22
Low	7.9	9.3	4.27	3.51

siderably more variable than the changes in latency. This is not surprising, since the amount of movement was an estimate made by *S* after the stimulus was removed, and so was subject to errors of *S*'s memory and interpretation, while latency was measured directly by *E*.

The analyses of variance also, of course, included various methodological variables. Variance due to *S*s was very highly significant in all cases. No effects of sequences of presentation could be found, probably because of the obscuring effect of the large subject variances. For the circles, which were presented first, the order of presentation (i.e., being the first, second, etc. stimulus to be presented) made a very highly significant difference, but not for the triangles. This is not surprising, since the so-called "first" triangular stimulus was actually the thirty-first stimulus presented. A separate analysis showed that the difference between circles and triangles produced a significant variance in amount of movement seen, but not in latency. The circles moved more than the triangles did.

Direction of movement.—A wide variety of directions of movement was seen by the successful *S*s. Even during the slightly more than 5 sec. of movement for each stimulus, some of the movements were curves rather than straight lines. However, it was possible in almost all cases to determine a direction in the frontal-parallel plane which fairly well represented each trial. In a few cases, two or three directions were reported for one trial, and these were counted separately. Then these directions were coded into four categories: 45°–135°, 135°–225°, 225°–315°, 315°–45°. Figure 2 shows the number of reports in each of these four directions, plotted

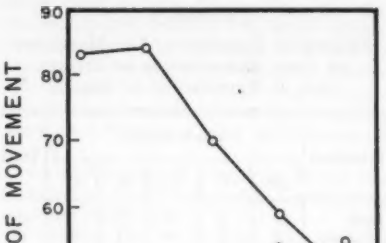

FIG. 2. Number of reports of movement in four directions as a function of size of stimulus

against size of stimulus. Since there were no differences between the directions seen with circles and triangles, they are combined in this figure and in the following discussion. It is at once apparent that the upward direction was most frequently seen, as is usual in studies of autokinetic movement (**3, 4**). A much more interesting result is that the number of times movement was seen in the upward direction decreased as the stimulus got bigger. Analysis of variance shows the interaction between size and direction to be significant. A possible reason for this is that when *S*'s face was blacked and he was observing the stimulus, the most visible part of his environment other than the stimulus was his own eyebrows, which were

least effectively concealed by the make-up. Consequently, as the size and brightness increased and, therefore, the illumination of the eyebrows got stronger, they provided a "frame of reference," and the chances of seeing movement in that direction were reduced.

Although the procedures for reporting direction of movement in this experiment were such that *S* had to report movement only in the frontal-parallel plane, several *Ss* nevertheless spontaneously reported seeing movement in the third dimension, sometimes combined with movement in the frontal-parallel plane and sometimes not. Seven of the 30 successful *Ss* gave these spontaneous reports on 30 out of 900 trials on which movement was seen. Four out of 21 rejected *Ss* also gave spontaneous reports of movement in the third dimension. These reports of three-dimensional movement were completely unexpected. It seemed likely that instructions designed to favor three-dimensional movement, and a procedure for reporting direction of movement which allowed for the possibility that three-dimensional movement might occur, would produce a good many more such reports. Therefore, an auxiliary experiment was designed.

Experiment II

Procedure

The apparatus and general procedure were the same as before. Five male *Ss*, naive about autokinetic movement, were used. Only the circles were used as stimuli, and all *Ss* saw the stimuli in the same order. If *S* failed to see movement on a stimulus, it was presented again later. No *Ss* were rejected. The *Ss* were instructed that they would be shown lights moving in three dimensions and that failure to see them as moving in three dimensions (i.e., seeing only two-dimensional movement) would be evidence of neuroticism. They were required to report the amount and direction of movement in the frontal-parallel plane, and in addition to report whether the stimulus came toward them, or away from them, or neither, and how much.

Results

So far as the effects of size and brightness are concerned, the results of this experiment are essentially the same as those of the previous experiment, and will not be further discussed. Table 3 shows the number of reports of various kinds of movement. (Note that the rows of the table do not add up to a fixed total. This is because if an *S* failed to see movement on one trial, the same stimulus was presented again later; if he failed a second time, it was not presented again. Only *S* 5 failed more than once on a given stimulus. This same problem arises again in Table 4.) It is apparent that the change in instructions and in reporting procedure resulted in a great increase in the number of reports of movement in the third dimension, an increase from less than 3% of all trials in the previous experiment to 28% of all trials in this experiment. No relation could be found between size or brightness of stimulus and number of reports of three-dimensional movement.

TABLE 3

Number of Reports of Various Kinds of Movement in Exp. II

S	In Frontal-Parallel Plane Only	In Third Dimension Only		Oblique Movements		Failures to See Movement
		Toward	Away	Toward	Away	
1	6	2	1	3	3	0
2	12	0	0	2	1	1
3	13	0	1	0	1	5
4	7	3	3	1	1	0
5	7	1	1	0	0	12
Total	45	6	6	6	6	18

Experiment III

Procedure

One criticism which could be made of the two previous experiments is that the instructions were so strongly worded that *S* might very well have lied, to avoid being thought neurotic. It is customary in experiments on autokinetic movement to use strong instructions in order to produce as many reports of movement as possible. However, since the argument from these experiments is that *S*s did see three-dimensional movement of large, bright stimuli, it is desirable to make sure that they were not lying. So another experiment was designed. It was identical in every respect with Exp. II, except for the instructions to *S*s. The *S*s were told: "I am going to show you several lighted objects. There are three groups of subjects in this experiment. For one group, all of these lighted objects will move. For another group, some of them will move and some of them will remain stationary. For the third group, all of the objects will remain stationary. I cannot tell you which group you are in. . . . (These objects) may move in any direction, up, down, right, left, toward you, away from you, or any combination of these directions . . ." Neuroticism was not mentioned.

Results

The results are just about as expected. Table 4 shows the number of responses in various categories. The most significant effect is a substantial increase in failures to see movement. Table 5 shows how these failures are distributed by size and brightness of stimulus. It is very clear that *S*s are more likely to fail to see large stimuli as moving. The distribution of failures among sizes is significantly different from chance beyond the .001 level. The effects of brightness are less clear. However, these results should not be interpreted to mean that the large stimuli are not seen as moving. The biggest stimulus was seen as moving 11 times out of 22 presentations, even with these very permissive instructions.

TABLE 4

Number of Reports of Various Kinds of Movement in Exp. III

S	In Frontal-Parallel Plane Only	In Third Dimension Only	Oblique Movements	Failures to See Movement
1	12	0	3	3
2	6	7	1	4
3	8	1	0	12
4	12	0	0	6
5	9	4	0	8
Total	47	12	4	33

TABLE 5

Number of Failures to See Movement on Various Stimuli in Exp. III

Brightness	Size in Inches					Total
	.25	1	4	8	12	
Low	0	0	0	4	4	8
Medium	0	2	3	5	3	13
High	2	2	2	2	4	12
Total	2	4	5	11	11	33

Discussion

This experiment has produced three main findings. The most important of them is that many *S*s do see large, bright autokinetic stimuli as moving, whether instructions are threatening or permissive. In Exp. II, the largest, brightest stimulus, which subtends about $7\frac{1}{2}°$, was seen to move by four *S*s and not seen to move by one. In Exp. III, it was seen to move by three *S*s, and not seen to move by two. (The figures in Table 5 are a bit misleading because if *S* failed to see movement once, he was shown the same stimulus again later, and usually failed to see it move the second time also, whereas successful stimuli were not repeated.) In Exp. I, of course, all successful *S*s saw the largest, brightest stimulus as moving.

The evidence also shows that those *S*s who do not see movement of large, bright stimuli also fail to see movement of other stimuli. Analysis of the data from the 21 rejected *S*s in Exp. I shows that they

failed on all stimuli fairly often. They failed more often on the larger, brighter stimuli. This would be expected, because the increase in latency with size and brightness shown by all *S*s would in the case of bad autokinetic *S*s bring their latencies over the arbitrary 60-sec. cutoff more often for the larger, brighter stimuli than for the smaller, dimmer ones.

The data suggest, though they do not prove, that the distribution of the population with respect to ease of seeing autokinetic movement is bimodal; this means that there are good autokinetic *S*s and bad autokinetic *S*s. If this is right, then the main effect of the rejection procedure in Exp. I was to screen out the bad autokinetic *S*s. If this distribution is not bimodal, then it is badly skewed, and the effect of the rejection procedure was to screen out the long tail of the distribution.

The second main finding of this experiment is that the latency of movement increases and the amount decreases as size and brightness increase. (This finding is consistent with the increased number of failures among rejected *S*s as size and brightness increase, as explained above.) Why is this true? The finding that the number of reports of movement upwards decreases as size and brightness increase suggests an answer. The black make-up applied to *S*'s nose and eyebrows did not completely remove them from the visual field. Of the two, the eyebrows were much more visible in this experiment. Naturally the visibility of these objects increased as size and brightness and therefore the amount of light illuminating the face increased. Probably these objects provided a visual "frame of reference" and so decreased amount and increased latency. This theory also suggests an explanation for the fact that size was more effective than brightness in hindering movement. The effect of increases in both size and brightness is to increase the amount of light falling on the face; but in addition the increase in size brings the contours of the autokinetic stimulus closer to the contours of the eyebrows, nose, and cheeks. If it can be assumed that the closer objects are the more likely they are to interact, it follows that increased size has two stabilizing effects, increased illumination and increased proximity to other contours, while increased brightness has only one. If some apparatus could be devised which would completely eliminate even the nose and eyebrows from the visual field, possibly no decrement in autokinetic movement with increased size and brightness of stimulus would be found.

The third main result of this experiment is that some autokinetic movement is three-dimensional. This occurs even with very permissive instructions, but more three-dimensional movements are seen if instructions are designed to encourage their occurrence. These movements may be straight toward or away from *S*, or may be combined with lateral or vertical movement. In most cases they occur some time (5 sec. or more) after the onset of the stimulus, and continue as long as the stimulus is on.

The results of this experiment are in apparent conflict with those of Karwoski, Redner, and Wood (**7**) and Luchins (**8, 9**). Both of those experiments reported no autokinetic movement for stimuli considerably smaller and dimmer than the biggest and brightest in this experiment (except that the Karwoski, Redner, and Wood stimulus was bigger in linear dimensions, though not in area, than the biggest in this experiment). The difference is probably attributable to the reduction screens and make-up used to eliminate illumination of irrelevant objects in this experiment.

In 1945 Graybiel and Clark (**5**) suggested that the real problem about autokinetic movement is not why the autokinetic stimulus moves but rather why the normal visual world does not. The results of this experiment reinforce this view. It seems at least possible that there is in principle no limit to the size and brightness of stimuli which will show autokinetic movement. What is so different about normal visual objects? The most obvious difference is that the visual world normally contains many objects, not just one. However, it is well known that collections of objects show auto-

kinetic movement, moving both as a unit and with relation to one another (though the sizes of the latter kind of movement are usually small). How diverse, bright, and large must a visual field be before it is stable? Another possibility to explore is that the visual world is full of complex shading effects, while autokinetic stimuli usually are not. Still another possibility is that the normal visual field is full of contours in the periphery as well as in the center. All of these questions lead to fairly straightforward experiments.

The usual notion for explaining the vagaries of autokinetic movement is the concept of "frame of reference." This phrase has a good deal of intuitive appeal, but its meaning is extraordinarily difficult to state precisely. How much of the visual environment must be removed before the frame of reference is gone? Can there be such a thing as a partial frame of reference? Is the frame of reference supplied by visual objects the same sort of thing as the social frame of reference that Sherif (10, 11) talks about? If not why are they different and if so why are they the same? Until the concept of frame of reference is sufficiently precisely specified so that questions like these are answerable, it will be difficult to use it effectively to account for phenomena of autokinetic movement other than those which led to its original formulation.

Summary

Autokinetic movement is usually thought of as exhibited by small, dim spots of light in a dark room. One reason for this is that it is difficult to build an apparatus in which large, bright spots of light do not illuminate other objects in the room. An apparatus was designed to solve this problem, and *S*s were required to report the latency, amount, and direction of autokinetic movement shown by spots of various sizes and brightnesses. The main finding was that stimuli as large as 7½° and as bright as 174 ft.-L. showed good autokinetic movement. The latency of movement increased and the amount decreased as size and to a lesser degree as brightness increased.

Some *S*s spontaneously reported movement in the third dimension, and so an auxiliary experiment was designed in which *S*s were given instructions specially favoring such movement. Under these circumstances, *S*s reported movement in the third dimension on 28% of all stimuli presented. No relationship between size or brightness and occurrence of three-dimensional movement could be found.

Another auxiliary experiment was designed to find out if *S*s were lying in the previous ones. Very permissive instructions were used. Fewer stimuli were seen as moving, but otherwise the results were the same as in the first two experiments.

In interpreting these results, the question of why the normal visual world does not show autokinetic movement is raised but not answered, and the concept of frame of reference is examined critically.

References

1. Adams, H. F. Autokinetic sensations. *Psychol. Monogr.*, 1912, **14**, No. 2 (Whole No. 59).
2. Bourdon, B. *La perception visuelle de l'espace.* Paris: Schleicher Frères, 1902.
2. Crutchfield, R. S., & Edwards, W. The effect of a fixated figure on autokinetic movement. *J. exp. Psychol.*, 1949, **39**, 561–568.
4. Edwards, W., & Crutchfield, R. S. Differential reduction of autokinetic movement by a fixated figure. *J. exp. Psychol.*, 1951, **42**, 25–31.
5. Graybiel, A., & Clark, B. The autokinetic illusion and its significance in night flying. *J. aviat. Med.*, 1945, **16**, 111–151.
6. Iwawaki, S. An experimental study of autokinetic movement. *Jap. J. Psychol.*, 1952, **23**, 203.
7. Karwoski, T. F., Redner, H., & Wood, W. O. Autokinetic movement of large stimuli. *J. gen. Psychol.*, 1948, **39**, 29–37.
8. Luchins, A. S. The autokinetic effect and gradations of illumination of the visual field. *J. gen. Psychol.*, 1954, **50**, 29–37.
9. Luchins, A. S. The autokinetic effect in central and peripheral vision. *J. gen. Psychol.*, 1954, **50**, 39–44.
10. Sherif, M. A study of some social factors in perception. *Arch. Psychol.*, 1935, No. 187.
11. Sherif, M. *The psychology of social norms.* New York: Harper, 1936.

(Received for early publication April 12, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

DIFFERENTIAL CONDITIONING AND LEVEL OF ANXIETY[1]

KENNETH W. SPENCE AND ROBERT S. BEECROFT

State University of Iowa

A series of experiments has been conducted recently in the Iowa laboratory that has been concerned with the relation of drive strength (D) to the level of response in conditioning (**5, 8, 9, 10**), extinction (**6**), and differential conditioning (**7**). Two methods of attempting to vary drive strength have been used: (*a*) direct variation of the intensity of the UCS employed, and (*b*) an indirect response-inferred variation on the basis of a score on a test of manifest anxiety (**11**). In connection with the latter method it was assumed that scores at the extremes (upper and lower 20%) of the anxiety scale would reflect differences in an individual's characteristic level of emotional responsiveness to his environment, particularly to any environmental situation that contained a threatening or noxious stimulus of any kind. Identifying, further, a high level of such emotional responsiveness with a high level of D, it was predicted on the basis of Hull's theory of $R = f(H \times D)$ that in simple, noncompetitional learning emotionally responsive *Ss* would exhibit a higher level of response (number of CR's) than emotionally nonresponsive *Ss*. The results of the studies on conditioning and extinction have supported this prediction, high anxiety groups consistently responding at a higher level than groups from the low end of the scale.

[1] This study was carried out as part of a project concerned with the influence of motivation on performance in learning under Contract N9onr-93802, Project NR 154-107 between the State University of Iowa and the Office of Naval Research.

In the case of differential conditioning the prediction of the theory is that high anxious (high D) *Ss* should respond at a higher level than low anxious (low D) *Ss* to both S^+ (reinforced stimulus) and S^- (nonreinforced stimulus). Furthermore, the difference between the excitatory strengths of S^+ and S^- should be greater for the high D than for the low D *Ss*. This derivation may be shown as follows, where H represents the habit strength of the positive stimulus and $\bar{H}$ represents the generalized habit strength of the negative stimulus:

$$E^+ = H \times D$$
$$E^- = \bar{H} \times D$$
$$E^+ - E^- = D(H - \bar{H})$$

Testing these derivations in a situation in which differential reinforcement was given to two stimuli, 500- and 5000-cycle tones, from the beginning of training, Spence and Farber (**7**) obtained data in two separate experiments that were in the predicted direction. However, the only differences that were significant were between the strength of response of the two groups to the positive CS.

While not statistically significant, the degree of differentiation (difference between E^+ and E^-) was consistently greater for the high-anxious *Ss* in four different instances. In contrast to this finding, Hilgard, Jones, and Kaplan (**2**) obtained a difference in the opposite direction, the nonanxious *Ss* showing the greater differentiation. The latter study differed from the Iowa experiment in that conditioning trials were given prior to

the introduction of the differential conditioning. While high-anxious *Ss* gave more CR's to the CS during this initial conditioning period than did low-anxious *Ss*, the difference was not significant as it has been in all of our studies.

The present investigation was conducted in order to test further the deductions from our theory that differential conditioning should be better in *Ss* scoring at the high end of the manifest anxiety scale than in *Ss* selected from the low end. The procedure of introducing the differential conditioning after conditioning to the positive CS had been established was employed. This procedure permitted a more direct comparison with the findings of Hilgard, Jones, and Kaplan. However, the experiment was completed in a single experimental session instead of two as it has been found that *Ss* are then much less likely to guess the purpose of the experiment than when they return for a second time.

Method

Subjects.—The *Ss* were selected from two courses in introductory psychology on the basis of scores made on a forced choice form of the Anxiety Scale (1). The anxious group consisted of 15 men whose scores on the test were in the upper 20% of scores made by Iowa students, while the nonanxious group was composed of 30 men whose scores were in the lower 20%. One other *S* was run in the experiment but was eliminated because of a power failure during conditioning. Since no *Ss* met the criterion employed in previous studies to define voluntary responders (8), none had to be eliminated on this basis.

Apparatus and method of recording.—The equipment for recording eyelid closure and presenting the UCS was identical with that used in a previous study (5). The positive CS (S^+) was a 500-cycle tone produced by a loudspeaker driven by a Hewlett-Packard oscillator. The negative CS (S^-), produced by the same means, was a 5000-cycle tone. The sound pressure levels of the tones at the position of *S*'s head, as measured by means of a General Radio sound level meter, Type 759-B, were 66 db. for S^+ and 49 db. for S^-. The duration of both tones was 615 msec., with the UCS occurring 490 msec. after the onset in the case of the positive CS. The UCS was omitted on trials on which the negative CS was presented. The duration of the UCS, an air puff of 1.0 lb./sq. in. applied to the right eye, was limited to 50 msec. by means of a 110-v., 60-cycle, AC-operated solenoid valve controlled by an electronic timer.

A CR was recorded whenever the record showed a deflection of 1 mm. or more in the interval 150 to 490 msec. following the onset of the CS. Responses with a latency of less than 150 msec. were classified as original responses to the CS and were not included in the data.

Conditioning procedure.—Following the reading of instructions, each *S* received three presentations of the positive CS alone. A single presentation of the UCS was then given. Immediately following these preliminary trials each *S* received 50 conditioning trials on which the positive CS was presented. On Trials 51–100 the negative CS was given 25 times and the positive CS 25 times. The order of presentation of S^+ and S^- was prearranged so that within each block of ten trials, five trials would be with S^+ and five trials with S^-, subject to the restriction that no more than two reinforced nor two nonreinforced trials occurred in succession. Intertrial intervals of 15, 20, and 25 sec., averaging 20 sec., and arranged according to a fixed schedule, were used throughout the 100 trials. At the end of the experiment *Ss* were warned not to discuss the experiment with other members of the class.

Results

Anxiety level and conditioning.—Figure 1 presents curves of acquisition of the CR in terms of percentage of responses occurring in successive ten-trial blocks and in the first two trials. As in the case of previous studies from this laboratory, the high-anxious *Ss* showed a higher level of response than did the low-anxious. In order to determine the significance of the difference in over-all level of response, the excitatory potential (E) values for each *S* for each block of ten conditioning trials were ascertained from Table 9 in Hull's *Principles of Behavior* (3) and a trend analysis of variance of these values was made. The results of this analysis are shown in Table 1.

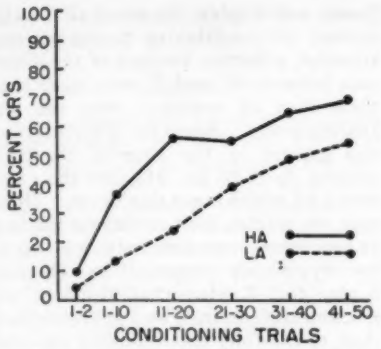

FIG. 1. Percentage of CR's on conditioning Trials 1–50 for high-anxious and low-anxious *Ss*

The *F* for the anxiety factor was found to be significant at better than the .01 level of confidence.

Table 2 presents the results for the last 20 trials (31–50) of conditioning in terms of the means and *SD*'s of the excitatory strengths. The difference between the means of the two groups for this period just preceding differential conditioning provided a *t* value of 1.81, which for 43 *df* is significant between only the .05 to .10 level on a double-wing hypothesis. Thus we see that while the difference between anxious and nonanxious *Ss* was highly significant over the whole 50 trials, it was not nearly so significant during the 20 trials just preceding the differential training.

Anxiety level and discrimination.—Figure 2 shows the level of performance of the two groups of *Ss* in terms of the percentage of CR's to the two stimuli (S^+ and S^-) during the 50 trials of differential conditioning, in which each stimulus was presented 25 times. As may be seen, the high-anxious *Ss* responded consistently at a higher level on both the positive and the negative stimulus although the difference was not very great in the case of the latter. The differences between the curves for the positive and negative stimuli were also larger in the case of the high anxious group. For purposes of statistical evaluation of the results of this period, the *E* values corresponding to the percentage of CR's made to each stimulus on the 25 trials were determined for each *S*. Table 3 presents the means and

TABLE 1

ANALYSIS OF VARIANCE OF *E* VALUES BASED ON TRIALS 1–50 OF HIGH-ANXIOUS AND LOW-ANXIOUS *Ss*

Source	*df*	*MS*	*F*	*p*
Anxiety	1	47.6028	7.54	<.01
Error (*b*)	43	6.3054		
Blocks	4	20.0350	34.90	<.001
Blocks × Anxiety	4	.9510	1.65	>.05
Error (*w*)	172	.5740		

TABLE 2

EXCITATORY STRENGTHS OF POSITIVE STIMULI ON TRIALS 31–50 FOR HIGH-ANXIOUS AND LOW-ANXIOUS *Ss*

Level of Anxiety	*N*	*E*+	
		Mean	*SD*
High	15	3.28	.92
Low	30	2.52	1.45

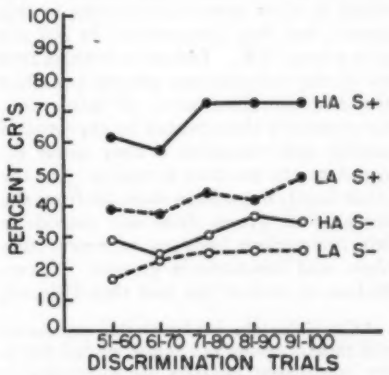

FIG. 2. Percentage of CR's to the positive and negative stimuli on discrimination Trials 51–100 for high-anxious and low-anxious *Ss*

TABLE 3

EXCITATORY STRENGTHS OF POSITIVE AND NEGATIVE STIMULI ON TRIALS 51–100 FOR HIGH-ANXIOUS AND LOW-ANXIOUS *Ss*

Level of Anxiety	*N*	E^+		E^-		$E^+ - E^-$	
		Mean	*SD*	Mean	*SD*	Mean	*SD*
High	15	3.11	.82	1.84	.85	1.26	.95
Low	30	2.15	1.41	1.37	1.24	.78	.73
t		2.37		1.30		1.82	
p		<.05		>.10		<.10	

SD's of the E^+, E^-, and $E^+ - E^-$ values for the high-anxious and low-anxious groups. In each instance it will be seen that the means are larger for the high-anxious group. Only in the case of the E^+ values, however, was the difference statistically significant (.02 level), although the difference between the discrimination measure ($E^+ - E^-$) was significant at about the .03 level if a single-wing hypothesis is employed.[2]

In the data presented above all *Ss*, whether or not they exhibited a conditioned response in the initial conditioning period, were included in the data. Actually, no *Ss* in the high-anxious group failed to show some conditioning in this period, but four low-anxious *Ss* did not give a single CR. Inclusion of these four *Ss* in the low-anxious groups penalizes the discrimination score of this group, for obviously they cannot be expected to exhibit differentiation if they never respond to the positive stimulus. On the other hand, to exclude these *Ss* from the low-anxious group does not provide a fair comparison between representative high- and low-anxious groups. Nevertheless, in view of the fact that Hilgard, Jones, and Kaplan discarded all *Ss* who showed no conditioning to the positive stimulus, a further analysis of the difference between E^+ and E^- was made with these four *Ss* excluded from the low anxious group. Again the differentiation was greater in the case of the high anxious *Ss* (1.26 vs. .91), but the *t* was only 1.33, which is not significant. However, the results, with or without the four *Ss*, are even more definitely opposed to the hypothesis proposed by Hilgard, Jones, and Kaplan that "anxious *Ss*, reacting more to their own apprehensions than to carefully discriminated environmental objects and relationships, will exhibit less discrimination of positive and negative stimuli than will non-anxious *Ss*" (2, p. 94).

In our previous study of differential conditioning (7), attention was called to the unknown role of inhibition in the theoretical treatment of the difference between high- and low-drive *Ss* and the suggestion was made that a more clear-cut test of our theory would be provided by comparison of the differential strength of response to the generalized stimulus prior to the possible effects of inhibition. Table 4 presents some data on the degree of differentiation between the positive and negative stimuli based on the percentage of responses to S^- on the first trial it was presented and to S^+ on the preceding trial. As will be seen, in agreement with our theoretical interpretation and against that of Hilgard, Jones, and Kaplan, the degree of differentiation was superior in the case of the high-anxious *Ss*. Evaluation of the difference between the differences by the method described by McNemar (4,

TABLE 4

PERCENTAGE OF HIGH-ANXIOUS AND LOW-ANXIOUS *Ss* RESPONDING TO S^+ ON TRIAL 50 AND S^- ON TRIAL 51

Level of Anxiety	Trial 50 Last S^+	Trial 51 First S^-	Diff.
High	66.7%	13.3%	53.4%
Low	43.3%	16.6%	26.7%

[2] Combining the results of this experiment with those of Spence and Farber, we find that in five independent instances the discrimination was better in the case of the high-anxious group. Such a result would be expected by chance about 3 times in 100.

p. 81), however, indicated that it was not significant.

Summary

An experiment was conducted to provide further evidence on simple and differential eyelid conditioning in high-anxious and low-anxious *Ss*. Each *S* received 50 conditioning trials with a 500-cycle tone as the positive CS. On Trials 51–100 the positive CS and the negative CS, a 5000-cycle tone, were each presented 25 times in a prearranged order. The interval between the positive CS and the UCS, a 1.0 lb./sq. in. air puff, was 490 msec. The UCS was omitted following the negative CS.

Excitatory strength to the positive CS during both conditioning and differential conditioning was significantly greater for the high-anxious subjects. Excitatory strength to the negative CS was also greater for the high-anxious *Ss* and the high-anxious *Ss* showed better discrimination (i.e., a greater difference in excitatory strength between the positive and negative CS) than did the low-anxious *Ss*. Though these latter results were not statistically significant, the differences in differentiation agree with previous results from this laboratory. These findings are consistent with expectations based on Hull's theory of the relation of drive and habit and are clearly contrary to the hypothesis proposed by Hilgard, Jones, and Kaplan that high-anxious *Ss* will show less discrimination.

References

1. Heineman, C. E. A forced-choice form of the Taylor Anxiety Scale. *J. consult. Psychol.*, 1953, **17**, 447–455.
2. Hilgard, E. R., Jones, L. V., & Kaplan, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, **42**, 94–99.
3. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
4. McNemar, Q. *Psychological statistics.* New York: Wiley, 1949.
5. Spence, K. W. Learning and performance in eyelid conditioning as a function of intensity of the UCS. *J. exp. Psychol.*, 1953, **45**, 57–63.
6. Spence, K. W., & Farber, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 116–119.
7. Spence, K. W., & Farber, I. E. The relation of anxiety to differential eyelid conditioning. *J. exp. Psychol.*, 1954, **47**, 127–134.
8. Spence, K. W., & Taylor, J. A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.
9. Spence, K. W., & Taylor, J. A. The relation of conditioned response strength to anxiety in normal, neurotic, and psychotic subjects. *J. exp. Psychol.*, 1953, **45**, 265–272.
10. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
11. Taylor, J. A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.

(Received for early publication July 23, 1954)

Journal of Experimental Psychology
Vol. 48, No. 5, 1954

THE RELATION OF ELECTRIC SHOCK AND ANXIETY TO LEVEL OF PERFORMANCE IN EYELID CONDITIONING[1]

KENNETH W. SPENCE, I. E. FARBER, AND ELAINE TAYLOR

State University of Iowa

In experiments involving classical defense conditioning it has been shown that the intensity of the aversive stimulus (UCS) producing the response is a factor determining the level of conditioning performance (**6, 9, 12**). This variable has been interpreted in terms of S-R theory as affecting the level of drive (*D*) of *S*, the value of *D* being assumed to be higher the more intense the UCS. Drive level, in turn, is assumed to be one of the factors determining the strength of the conditioned response. According to this interpretation, the effects, presumably emotional, of the noxious UCS are assumed to be relatively persistent, extending well beyond the intertrial intervals usually employed in conditioning experiments. Thus the drive level (*D*) present at the moment of occurrence of a conditioned anticipatory response is conceived to be, in part at least, a function of the residual effects of the emotional reaction to the noxious stimulation of the preceding trials.

Also in line with this interpretation that drive strength (*D*) is a function of *S*'s emotional level are the investigations in which individuals, varying in emotional responsiveness as determined by a manifest anxiety scale (**14**), have been shown to exhibit different levels of performance in classical defense conditioning (**10, 11, 12, 13**). It has been assumed in these experiments that *Ss* scoring at the extremes (e.g., upper and lower 20%) of the anxiety scale exhibit characteristic differences in the magnitude or strength of their emotional reaction to the same intensity of UCS and thus have different levels of *D* operative.

That the motivational level of *S* may be raised by the administration of a noxious stimulus (shock) prior to measurement in a test situation has been demonstrated recently by Amsel and Maltzman (**3**) and Siegel and Brantley (**8**). These investigators found that the level of consummatory response (drinking, eating) in a test period immediately following the receipt of a shock was significantly increased. Another example in which the response that was increased in strength was a learned one is that provided by Nagaty (**5**). This investigator first conditioned albino rats to avoid a shock by rotating a wheel to the sound of a buzzer. Subsequently, *Ss* were divided into three groups and tested under different extinction conditions: (*a*) shock given 1 sec. before the buzzer; (*b*) shock given 20 sec. before the buzzer; and (*c*) no shock preceding the buzzer. All three groups showed a decrement in response strength during 50 test trials typical of extinction, but the performance curves of the two shocked groups were significantly above those of the group not given any shocks. These investigators all interpreted their results as reflecting an increase in the generalized drive level of their *Ss* resulting from persisting emotional effects. Subsequent experiments by

[1] This study was carried out as part of a project concerned with the influence of motivation on performance in learning under contract N9onr-93802, Project NR 154-107 between the State University of Iowa and the Office of Naval Research.

Amsel and others (1, 2) have shown that whether the addition of such an emotional drive to the motivational complex will lead to an increase in response strength depends upon the conditions of introduction of the noxious stimulus. Important among these is whether the emotion-arousing stimulus cues elicit overt responses that compete with the criterion response.

The present experiment was undertaken with a view to ascertaining the effect of heightening the emotional level of *Ss* upon the performance level in eyelid conditioning either by the administration of an electric shock occasionally during the course of conditioning or by the threat of giving such a stimulus without actually doing so. One group of *Ss* (II) was given a total of 11 shocks throughout 80 conditioning trials, while a second group (III) was given no shocks until the final trial, although prior to the conditioning the *Ss* were required to adjust an electric shock to a level they deemed uncomfortable and were led by the instructions to believe that they would occasionally receive this shock. Group I *Ss* never received a shock at any time and did not have the shock electrodes attached to them.

Method

Subjects.—A total of 74 women students from an introductory course in psychology served as *Ss*. Eight of these were eliminated because they met the criteria employed in previous studies in this laboratory defining a voluntary responder (12). Four *Ss* who misunderstood the instructions and two who gave CR's to pretest trials were also eliminated. The remaining 60 *Ss* were randomly assigned to three groups of 20 each: I, no shock; II, shock; and III, threat of shock.

Apparatus and method of recording.—The *S* was seated in a dental chair in a semi-soundproof, darkened room. The *E* was in an adjoining room where the stimulus controls and recording equipment were located. The equipment for recording the eyeblinks and presenting the UCS was identical with that used in a previous study (9).

The CS consisted in an increase in the brightness of a 6-cm. circular disc from a level of .05 to 1.51 apparent foot candles. The duration of the CS was 550 msec., with the UCS occurring 500 msec. after the onset of the CS. The duration of the UCS, an air puff of .6 lb./sq. in. applied to the right eye, was limited to 50 msec. by means of a 110-v., 60-cycle, AC-operated solenoid valve controlled by electronic timers.

The ready signal, consisting of a 1000-cycle tone of 50 msec. duration preceded each presentation of the CS by 3 sec. In the case of the shock group this auditory signal was accompanied by an electric shock on some of the trials. The shock was provided by the discharge of a 4 microfarad condenser and was administered through electrodes attached to the index and middle fingers of the right hand. The *Ss* were instructed to blink upon presentation of the ready signal and then to fixate the circular disc in front of them. A CR was recorded whenever the record showed a deflection of 1 mm. or more in the interval 200 to 500 msec. following the onset of the CS.

Procedure.—Following the reading of instructions, all *Ss* were given three preliminary trials to the CS alone. One preliminary presentation of the air puff was then given and *S*'s eyelid reactions recorded for 30 sec.

The level of shock employed was determined individually for each *S* of Groups II and III. The initial presentation of the shock was at a near threshold level and it was then increased in 10-v. steps until *S* reported it as "uncomfortable but not painful." The shock setting remained at this level throughout the period of conditioning.

Immediately following the preliminary trials and the shock adjustments, 80 pairings of the CS and UCS were given. The intertrial interval was 20, 25, or 30 sec., with the average interval being 25 sec. Shock was administered simultaneously with the ready signal on Trials 1, 9, 16, 25, 34, 42, 49, 56, 65, 73, and 80 for Group II and on Trial 80 only in the case of Group III.

Following Trials 20 and 50 a section of the instructions emphasizing the importance of remaining relaxed and not attempting to control the response of the eye, was reread to *S* over a microphone. At the completion of the experimental session each *S* was asked to refrain from discussing the experiment. Unfortunately, most of the *Ss* were from the College of Nursing and were housed together in the same dormitory. Questioning indicated that while details of the experiment were probably not discussed widely,

the fact that shock was used in some instances soon became common knowledge.

Results and Discussion

Effects of shock and shock threat.—Curves of acquisition of the CR based on the percentage of responses occurring in successive blocks of ten trials for the three groups are presented in Fig. 1. As may be seen the two shock-involved groups (II and III) exhibited a higher level of performance than the nonshock group (I) in the early portion of the conditioning (Trials 1–30). However, this advantage disappeared, there being little difference among the three groups in the latter half of the trials. In view of the obvious lack of difference between Groups II and III, their data were combined for purposes of statistical analysis. Table 1 presents the results of an analysis of variance involving number of CR's occurring on successive blocks of 20 trials. As may be seen, the over-all difference between the groups (no shock versus shock involvement) was not significant, but the interaction (groups by blocks of trials) was significant at the .05 level. This finding indicates that

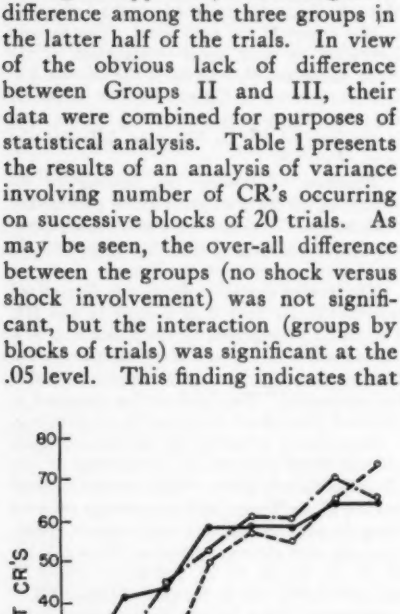

Fig. 1. Percentage of CR's for three groups of *Ss* in blocks of ten trials

TABLE 1

Analysis of Variance of Number of CR's in Four Blocks of 20 Trials

Source	*df*	*MS*	*F*	*p*
Groups	1	67.50		
Error	58	88.31		
Blocks	3	774.32	76.51	<.001
G × B	3	27.75	2.74	<.05
Error (*w*)	174	10.12		

the differences between the compared groups are a function of the stage of conditioning. A *t* test of the differences within each of the four blocks of trials indicated that the difference during Trials 1–20 was significant at about the .05 level and that during Trials 21–40 at about the .20 level. The other differences were negligible.

These findings suggest that in the case of the shock group there was an adaption to the shock, which it will be recalled, was adjusted initially only to a value that was "uncomfortable but not painful." The general report of the *Ss* agreed with this interpretation, most *Ss* indicating that the shock seemed to become milder. The fact that the shock-threat group also did not show a higher level of performance than the nonshock group at the later stages of training suggests that these *Ss* lost some of their fear of the threatened shock as they failed to receive any more shocks.

Interrelation of shock and anxiety.—Since scores on the Taylor Manifest Anxiety Scale were available for all *Ss* it was possible to study the interrelation of this factor with shock effects. The *Ss* were divided into two groups, those who scored above 14 on the anxiety scale being designated as high anxious and those 14 or less as low anxious.[3] Table 2 presents

[3] Attention is directed to the point that *Ss* were randomly selected so far as the anxiety scale was concerned, and the subsequent division roughly into the upper and lower 50% does not

the means and *SD*'s of the number of CR's made in the first half of conditioning (Trials 1–40). It will be observed that the nonshock group was divided evenly into high- and low-anxiety subgroups of ten *S*s each, but that in the case of the shock-involved group there were many more low- than high-anxious *S*s (25 versus 15).

Further examination of Table 2 shows that the difference between the nonshock and shock conditions was much greater in the case of the high-anxious *S*s than in the low-anxious *S*s. Whereas the latter difference was not significant, the difference (22.60 vs. 12.90) in the case of the high-anxious

TABLE 2

NUMBER OF CR'S OCCURRING ON TRIALS 1–40

Condition	Low Anxious			High Anxious		
	N	Mean	*SD*	*N*	Mean	*SD*
No shock (I)	10	9.90	8.18	10	12.90	8.09
Shock (II & III)	25	12.04	9.53	15	22.60	7.62

*S*s was significant at better than the .01 level ($t = 2.92$). Comparing the low- and high-anxious *S*s it will be observed that in both sets of conditions the high-anxious *S*s gave a larger number of CR's. Only in the case of the shock condition, however, was the difference significant and again it was beyond the .01 level of confidence ($t = 3.55$).

In order to run an analysis of variance (factorial design) of the data in Table 2, the last ten *S*s run in the shock-low anxious category were discarded, thus equating the *N* in both high and low anxious groups under the shock conditions at 15 *S*s each. Table 3 presents the results of analysis of variance of these data. The *F*

represent as widely separated groups on this scale as have been used in our previous studies.

TABLE 3

ANALYSIS OF VARIANCE OF NUMBER OF CR'S IN TRIALS 1–40

Source	*df*	*MS*	*F*	*p*
Groups	1	371.85	5.38	<.05
Anxiety	1	792.02	11.45	<.01
G × A	1	205.01	2.96	>.05
Within cells	46	69.15		
Total	49			

for the groups (nonshock vs. shock conditions) was significant at the .05 level, while the *F* for anxiety was significant at better than the .01 level. Somewhat surprisingly, the interaction between groups and anxiety was not significant.

These findings have a bearing on two alternative hypotheses that we have entertained concerning the interpretation of high anxiety. One is that *S*s at the high end of the scale are chronically emotional and have a higher level of *D* at all times than do *S*s from the low end of the scale (14). A second, somewhat different conception is that *S*s from the higher end of the scale react more emotionally to noxious forms of stimulation or threatening situations than do *S*s from the low end of the scale and that this emotional response adapts more slowly.[3] According to the first of these views, high anxious *S*s would be expected to exhibit a higher level of conditioning than low anxious *S*s under all conditions, even, for example, in classical reward conditioning in which there is no threatening or noxious stimulus. The second

[3] This interpretation is similar to that proposed some years ago by Rosenbaum (7), who found that anxious *S*s showed a higher gradient of stimulus generalization than nonanxious *S*s only when the experimental conditions involved the administration of a relatively strong electric shock. It is important to note that, although it bears a superficial resemblance to the suggestion of Hilgard, Jones, and Kaplan (4) that the effects of anxiety are somehow related to threat, the present interpretation continues to be founded on the assumption that high emotional responsiveness or anxiety reflects a high level of *D* and low anxiety a relatively low level of *D*.

conception would predict that the difference between high and low anxious *S*s in classical conditioning situations would be a function of the degree of noxiousness or threat in the experimental situation. The data of the present experiment, it will be noted, are more in accord with the second conception. When there was no shock or threat of shock and only a relatively mild air puff as the UCS, the difference between the high and low anxious *S*s was small and not significant. However, when shock or threat of shock was introduced, the difference between high and low anxious *S*s was large and highly significant.

Summary

In order to ascertain the effect of heightened emotional level upon performance in eyelid conditioning, three groups of 20 women *S*s each were given 80 conditioning trials. One group was given an occasional electric shock in conjunction with the ready signal preceding the CS; another group was told it might receive such shocks, but actually received them only during a preliminary tryout; and the control group received neither shock nor the threat of shock.

The results showed no difference between the shocked *S*s and those that received the threat of shock. When these two "shock involved" groups were combined, they showed a reliably higher level of performance than the control group during the earlier stages of conditioning.

When each group was divided into upper and lower levels in terms of scores on the Taylor *A* Scale, the relatively anxious *S*s gave more CR's during the first half of conditioning than did the relatively nonanxious *S*s. However, the difference between anxiety levels was significant in the case of the shock-involved *S*s only.

On the basis of these results, it was concluded that level of emotionality, as defined by the presence or absence of shock or threat of shock, is related to performance in eyelid conditioning. And, the effect of manifest anxiety upon this performance may be a function of the degree of noxiousness or threat in the experimental situation.

References

1. Amsel, A. The effect upon level of consummatory response of the addition of anxiety to a motivational complex. *J. exp. Psychol.*, 1950, **40**, 709–715.
2. Amsel, A., & Cole, K. F. Generalization of fear-motivated interference with water intake. *J. exp. Psychol.*, 1953, **46**, 243–247.
3. Amsel, A., & Maltzman, I. The effect upon generalized drive strength of emotionality as inferred from the level of consummatory response. *J. exp. Psychol.*, 1950, **40**, 563–569.
4. Hilgard, E. R., Jones, L. V., & Kaplan, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, **42**, 94–99.
5. Nagaty, M. O. The effect of reinforcement on closely following S-R connections: I. The effect of a backward conditioning procedure on the extinction of conditioned avoidance. *J. exp. Psychol.*, 1951, **42**, 239–246.
6. Passey, G. E. The influence of intensity of unconditioned stimulus upon acquisition of a conditioned response. *J. exp. Psychol.*, 1948, **38**, 420–428.
7. Rosenbaum, G. Stimulus generalization as a function of clinical and experimentally induced anxiety. Unpublished doctor's dissertation, State Univ. of Iowa, 1950.
8. Siegel, P. S., & Brantley, J. J. The relationship of emotionality to the consummatory response of eating. *J. exp. Psychol.*, 1951, **42**, 304–306.
9. Spence, K. W. Learning and performance in eyelid conditioning as a function of intensity of the UCS. *J. exp. Psychol.*, 1953, **45**, 57–63.
10. Spence, K. W., & Beecroft, R. S. Differential conditioning and level of anxiety. *J. exp. Psychol.*, 1954, **48**, 399–403.
11. Spence, K. W., & Farber, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 116–119.
12. Spence, K. W., & Taylor, J. A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.
13. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
14. Taylor, J. A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.

(Received for early publication July 23, 1954)